Matleena Myntti

Target identification with hyperspectral lidar

School of Science

Thesis submitted for examination for the degree of Master of Science in Technology.

Espoo 20.3.2015

Thesis supervisor:

Prof. Matti Kaivola

Thesis advisor:

D.Sc. (Tech.) Albert Manninen and M.Sc. (Tech.) Teemu Kääriäinen

Aalto University
School of Science

AALTO UNIVERSITY
SCHOOL OF SCIENCE

ABSTRACT OF THE
MASTER'S THESIS

Author: Matleena Myntti

Title: Target identification with hyperspectral lidar

Date: 20.3.2015 Language: English Number of pages: 6+85

Department of Applied Physics

Professorship: Optics and photonics Code: Tfy-125

Supervisor: Prof. Matti Kaivola

Advisor: D.Sc. (Tech.) Albert Manninen and M.Sc. (Tech.) Teemu Kääriäinen

Hyperspectral lidar is a remote sensing technique based on active broadband illumination of a target and measuring a spectrum of the backscattered light. Hyperspectral imaging that uses ambient lighting (such as sunlight) is widely used in numerous applications ranging from crop management to military applications, yet actively illuminating long-distance lidar devices have been lacking until recently. A fully automated long-range hyperspectral imaging lidar instrument was built and tested in the Centre for Metrology and Accreditation (Mittatekniikan keskus, MIKES) in 2013. This device could be used for spectral imaging for the wavelength band 1000–2300 nm. However, the control software lacked spectral analysis and material identification functionality.

This Master's thesis introduces the implementation and testing of a solid material identification software for the hyperspectral lidar imaging instrument. The software is based on five different spectral matching algorithms, which are the Euclidean distance, the spectral angle mapper (SAM), the linear correlation, the multidimensional normal distribution (MND), and the spectral information divergence (SID) algorithms. These algorithms use several library spectra of each material type in order to express the spectral variability of materials; this differs from the usual approach in the literature, in which the measured spectra are normally compared against one library spectrum per each material type candidate. The methods were tested with different band divisions, variance scaling parameters, and material and library spectrum classification choices.

The software shows promising potential, even though limitations in the lidar instrument bring challenges. In laboratory environment, the instrument is capable of accurate material recognition, while field measurements were proven more challenging. However, the instrument and the software have a few characteristics that can be improved, giving the setup potential to give even better performance.

Keywords: Lidar, Hyperspectral, Remote sensing, Supercontinuum, Target recognition, Target identification, Spectral identification software, Spectral matching

AALTO-YLIOPISTO
PERUSTIETEIDEN KORKEAKOULU

DIPLOMITYÖN
TIIVISTELMÄ

Tekijä: Matleena Myntti

Työn nimi: Kohteen tunnistus käyttäen hyperspektristä lidaria

Päivämäärä: 20.3.2015 Kieli: Englanti Sivumäärä: 6+85

Teknillisen fysiikan laitos

Professuuri: Optiikka ja fotoniikka Koodi: Tfy-125

Valvoja: Prof. Matti Kaivola

Ohjaaja: TkT Albert Manninen ja DI Teemu Kääriäinen

Hyperspektrinen lidar on kaukotunnistusmenetelmä, joka perustuu kohteen aktiiviseen valaisuun laajalla aallonpituuskaistalla sekä heijastuneen valon spektrin mittaamiseen. Ympäristön valoa kuten auringonvaloa hyödyntävät kaukokuvantamismenetelmät ovat olleet pitkään käytössä lukuisissa eri sovelluksissa aina viljapeltojen kuvantamisesta maanpuolustukseen. Sen sijaan aktiivista valaisua käyttäviä kaukokuvantavia lidar-laitteita ei ole juuri rakennettu ennen kuin aivan viime vuosina. Vuonna 2013 Mittatekniikan keskuksessa (MIKES) rakennettiin ja testattiin automatisoitu kaukokuvantava hyperspektrinen lidar-instrumentti, jota voidaan käyttää spektriseen kuvantamiseen 1000–2300 nm kaistalla. Laitteen ohjausohjelmistosta kuitenkin puuttui vielä toiminnot spektrien analysoinnille ja materiaalitunnistukselle.

Tässä diplomityössä esitellään kiinteiden aineiden tunnistusohjelmiston kehitys ja testaus rakennetulle hyperspektriselle lidar-kuvantamisinstrumentille. Ohjelmisto käyttää viittä eri spektristä sovitusalgoritmia, joiden perustana ovat euklidinen etäisyys, spektrinen kulmakartoittaja (SAM), lineaarinen korrelaatio, moniulotteinen normaalijakauma (MND) sekä spektrisen informaation divergenssi (SID). Nämä algoritmit käyttävät useita kirjastospektrejä per materiaali ilmaistaakseen spektristä vaihtelua, mikä poikkeaa tavanomaisesta lähestymistavasta eli mitatun spektrin vertailusta yhteen kirjastospektriin per materiaalivaihtoehto. Näitä metodeita kokeiltiin käyttäen erilaisia kaistajakoja, varianssiskaalausparametreja sekä materiaalityyppi- ja kirjastoluokitteluvalintoja.

Ohjelmistossa näyttää olevan lupaavia mahdollisuuksia, joskin lidar-instrumentin rajoitukset tuovat haasteita. Laboratorio-olosuhteissa instrumentti kykeni tarkkaan materiaalitunnistukseen, mutta kenttämittaukset osoittautuivat haasteellisiksi. Instrumentissa ja ohjelmistossa on ominaisuuksia joita voi yhä parantaa, ja siten laitteen toimintakykyä voidaan parantaa entisestään.

Avainsanat: Lidar, Hyperspektrinen, Kaukotunnistus, Superjatkumo, Kohteentunnistus, Spektrinen tunnistusohjelmisto, Spektrinen sovitus

Preface

I would like to thank my advisors D.Sc. (Tech.) Albert Manninen and M.Sc. (Tech.) Teemu Kääriäinen from MIKES Metrology at VTT Technical Research Centre of Finland Ltd. for guidance and support, as well as my superior D.Sc. (Tech.) Mikko Merimaa for all the help and my supervisor Professor Matti Kaivola from Aalto University School of Science for his keen interest towards my work and the meticulous proofreading.

I would also like to thank the Scientific Advisory Board for Defence (Maanpuolustuksen tieteellinen neuvottelukunta, MATINE) for funding this project and Scott Buchter and Miika Heiliö from Lasersec Systems Corporation for providing us a spectrometer. Additionally I would like to thank Tomi Parviainen from the Finnish Defence Research Agency (FDRA) and Sanna Kaasalainen, Teemu Hakala and Olli Nevalainen from Finnish Geospatial Research Institute (FGI) for their long-term co-operation with us.

I would like to give my gratitude to HR secretary Taina Sarvikas from MIKES for helping me in practical matters, to my upper secondary school physics teacher Sirpa-Liisa Rauta from Porin suomalaisen yhteislyseon lukio for encouraging me to start studying this demanding degree, and to studies secretary Taru Bister from Aalto University School of Science, whose help over these years has been priceless.

Finally, I would like to thank my parents, Erja and Jukka Myntti, for all their love and support.

Otaniemi, 31.3.2015 Matleena M. A. Myntti

Contents

1 Introduction

Electromagnetic radiation propagates within matter in a predictable manner, following well-defined physical laws. Information about the expected radiation behavior enables determination of the beam propagation interactions by analyzing its transmitted or back-scattered spectrum. This can be used to determine the properties of the medium, which offers the basic principle for various radiation-based measurement and imaging techniques, such as medical X-ray imaging. Electromagnetic radiation is the base of several modern remote-sensing applications, because this kind of radiation is easy to generate and it provides plenty of data upon interactions with matter. In addition, the wide spectrum of electromagnetic radiation can be used to acquire plenty of data from the target by using spectral analysis.

Optical remote sensing technologies can be divided into two main categories: passive remote sensing and active remote sensing. Passive remote sensing relies on ambient lighting, such as sunlight, and uses that for sensing the target. However, the effective use of passive sensing methods is limited by the availability of illumination; for example, solar-sensitive measurement instruments are less reliable during night or in shadowed areas. Meanwhile, active remote sensing uses an integrated radiation source with well-known properties. While active remote sensing techniques require accurate adjustment to collect the back-scattered or transmitted radiation beam as precisely as possible, it also offers more accurate and more reliable spectral data than their passive counterparts. [1]

Active remote sensing techniques that use optical electromagnetic radiation are alternatively known as lidar (LIght Detection And Ranging). Lidar employs one or multiple wavelengths from ultraviolet (UV, 100 nm) to far-infrared ranges (far-IR, 100 μm) depending on its intended application. Lidar is widely employed in various fields, ranging from simple distance and density measurements to atmosphere studies, crop management, and military detection applications. The main focus of this thesis, hyperspectral lidar, employs a wide spectrum of IR radiation rather than separate monochromatic wavelengths. The wide spectrum is often created by supercontinuum (SC) light generation. [2, 3]

Although passive hyperspectral remote sensing has been commercially used in hyperspectral cameras for years, developers of the field have built only few prototypes of hyperspectral lidar instruments. Moreover, most of these instruments have been costly due their use of expensive commercial SC light sources that also have only low-to-moderate power output. The limited power output prevents applying these SC sources for lidar devices intended for long-distance target recognition.

To answer the need of an active long-distance target recognition instrument, an automated long-distance hyperspectral lidar device has been developed at the Centre for Metrology and Accreditation (Mittatekniikan keskus, MIKES) in 2013 [4, 5]. The required high-power SC light source was built using multi-mode graded-index optical fiber and a high power pump laser source with a nominal wavelength of 1064 nm. This custom-made SC source provides 16 W of averaged power for a wavelength band of 1000–2300 nm, and the device has been demonstrated to be effective for long-distance (1.5 km) hyperspectral measurements [4]. In order

to apply the developed lidar instrument for actual target material recognition, the setup also requires software that would analyze the measured reflection spectra and identify the target material in real time.

The aim of this thesis is to evaluate the performance of five spectral identification methods in order to recommend the most effective methods for use in a lidar instrument developed by MIKES. The following five methods were chosen for this work: distance method, spectral angle mapper method (SAM), correlation method, multidimensional normal distribution (MND) method, and spectral information divergence method (SID) [2, 6]. The methods were chosen due to their simplicity and fast calculation speed as well as overall suitability for material identification simultaneously with scanning. As an unusual approach to most of them, these methods are here applied to material identification by using multiple reference spectra for each material rather than single averaged ones: this allows the expression of spectral variability among a single material type and also combining several spectra flexibly into supercategories. This kind of approach is common to the MND method but not to the others, as the others rather use a single reference spectrum for each material.

The identification methods are evaluated in terms of their reliability, calculation speed and ability to clearly differentiate the spectra of different materials. To accomplish this, the thesis identifies available algorithms for detecting material spectra and tests their application for the spectra measured by the lidar instrument. The identification procedure will be used as a part of the scanning function of the device.

The thesis uses the hyperspectral lidar instrument presented previously in Kääriäinen's work [4]. However, since the spectrum of the initial SC output contains high-energy short-wavelength peaks that can endanger users' or by-passers' eyesight, the setup is modified to filter out wavelengths below 1500 nm. Using this modification, the setup fulfills the eye safety requirements set by the Radiation and Nuclear Safety Authority (Säteilyturvakeskus, STUK) [7]. In addition, the identification methods use only the wavelength band of 1500–2200 nm, as the filtered output of the SC source is brightest in this wavelength band.

The structure of this work proceeds as follows. Chapter 2 explains the core mechanisms of hyperspectral lidar, detailing the physical phenomena employed in the spectral identification and requirements for a reliable identification process. Chapter 3 describes the hyperspectral lidar setup used with the identification algorithms based on the instrument built in 2013 [4]. Chapter 4 introduces the identification algorithms used in this thesis and the implementation of the methods into the identification software, the challenges faced by the identification procedure, and the presented solutions. The same chapter also introduces the tests used to evaluate the identification algorithms. Finally, Chapter 5 details the results of the algorithm comparison tests and the overall performance of the software, and Chapter 6 ties up this thesis into conclusions.

The project related to this master's thesis was funded by the Scientific Advisory Board for Defence (Maanpuolustuksen tieteellinen neuvottelukunta, MATINE), and the experiments were carried out in MIKES during the summer and fall in 2014. In 2015, the thesis finalized in MIKES Metrology at VTT Technical Research Centre of Finland Ltd., which MIKES merged with at the beginning of 2015.

2 Hyperspectral sensing

In hyperspectral lidar, the target is illuminated by a light beam of wide spectral bandwidth, and then the back-reflected spectrum is measured and analyzed. Comparison between the emitted and reflected spectra offers information about the properties of the target material and/or the atmosphere between the reflection target and the lidar instrument. In this thesis, the wide bandwidth is created by using a supercontinuum (SC) laser source. The following sections introduce the operational principles of hyperspectral sensing, starting from the electromagnetic radiation behavior on solid surface, then explaining the radiation scattering and atmosphere effects between the instrument and the target, and finally describing the rudimentary features of SC light generation and spectral detection.

2.1 Behavior of electromagnetic radiation on solid surface

Following the law of energy conservation, all electromagnetic radiation must proceed somewhere when the radiation confronts a surface. The three ways of radiation behavior at a surface are reflection, absorption and transmission, and their relation can be expressed as

$$R(\lambda) + A(\lambda) + T(\lambda) = 1, \tag{1}$$

where λ is the wavelength, R is the reflectance, A is the absorbance and T is the transmittance of the target material. Reflectance, absorbance and transmittance are material-specific properties that are defined as

$$R(\lambda) = \frac{I_r(\lambda)}{I_i(\lambda)}, \quad A(\lambda) = \frac{I_a(\lambda)}{I_i(\lambda)}, \quad T(\lambda) = \frac{I_t(\lambda)}{I_i(\lambda)}, \tag{2}$$

where I_i is the intensity of the incident radiation on the target and I_r, I_a, and I_t are the intensities of the radiation intensity reflected from, absorbed into, or trasmitted through the target surface, respectively. The ratios of these three quantities at different wavelengths depends on the properties of the target material. This makes different materials identifiable by spectral remote sensing techniques. In this work, target identification is based on the obtained reflection spectra. In addition to atomic composition, the spectrum is also affected by the shape of the target and the atmospheric effects between the target and the lidar instrument.

The next subchapter introduces the rudimentary basics of quantum-mechanical radiation behavior on solid surface; for more details on solid state physics and hyperspectral remote sensing, the reader is advised to refer to [2, 8]. After declaring the theory, we discuss the scattering mechanisms of electromagnetic radiation and the effect of atmosphere on measurements.

2.2 Brief quantum-mechanical explanation of absorption spectrum

The reflection spectra are based on the behavior of photons in matter. As seen by quantum physics, physical system can have energy states on only certain discrete energy levels, which depend on the potential and the kinetic energies of all particles in the system. For example, in Born-Oppenheimer approximation, the energy state of a molecule consists of the potential energies of electrons and vibrational and rotational energies of the nuclei, assuming the molecule is fixed into a position.

In order to change the energy state of the system, such as the potential energy of an electron or the rotational state of a molecule, the system must absorb or emit a photon with energy that equals the energy state difference. This energy state difference is also known as an energy gap. In solid matter, innumerable atoms and/or molecules are tightly connected to each other, which results in a network of minutely differing energy gaps. This results in a large variety of photon energies that the material can absorb, and therefore, solid materials have wide and smooth absorption bands rather than a peaked absorption spectrum. Therefore, instead of identifying certain material peaks, identification of solid materials relies on identifying the general shape of the spectrum.

The electronic transitions require more energy than vibrational transitions, and vibrational transitions require more energy than rotational transitions. Because of this, the electronic transitions occur at the UV range, and the vibrational and rotational transitions happen at the IR and far-IR ranges, respectively. [2]

2.3 Radiation scattering

In addition to atomic structure, radiation scattering also depends on the surface roughness, target shape, and the incident angle of the radiation. These properties add to the reflection spectrum characteristics that are not dependent on the material itself, which makes the characteristic material reflection spectrum less identifiable. Therefore, the geometric properties need to be taken into account in spectral analysis.

Electromagnetic radiation scattering can be defined as an interaction of photon with matter, which causes the photon to change the direction of its propagation [9]. The scattering can be divided into specular reflection and diffuse scattering, former one being dominant on very smooth surfaces and latter one on rough (diffusive) surfaces. All real-life scattering is a combination of specular reflection and diffuse scattering. The scattering types are illustrated in Figure 1.

These kinds of scattering can be modelled by bidirectional reflectance distribution function (BDRF). However, typically exact combinations of specular reflection and diffuse scattering are laborious to calculate, so it is beneficial to use approximations to model the scattering behavior. The simplest model is to assume that all back-scattered light is caused by diffuse scattering. In most cases, this model is convenient as long as the direction of the specular reflection can be avoided by tilting the target surface. [2, 10]

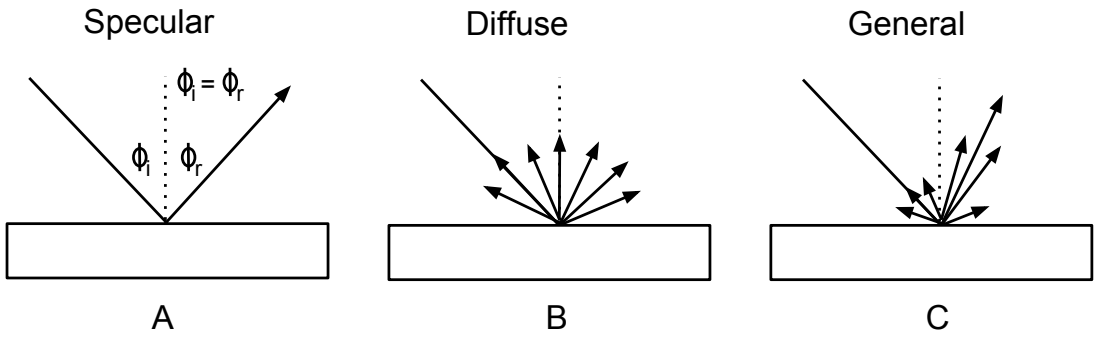


Figure 1: Visualization of light scattering on surface: Specular reflection (A) applies to perfectly smooth surfaces and is predictable by Snell's law. Diffuse scattering (B) applies to perfectly diffusive surfaces and assumes that the radiation is spread equally in every direction in a hemisphere. In real life, all scattering is a combination of these two mechanisms (C). [4]

2.4 Atmospheric effects

The radiation used in lidar does not interact only with the target material and the instrument, but it also interacts with the air between them. The atmospheric particles scatter and absorb the radiation, thus affecting its spectrum. The scattering process can be either elastic or inelastic; elastic scattering refers to the process in which the photon bounces from the scatterer with no kinetic energy changes, and inelastic scattering refers to photon being partially absorbed into the scatterer and therefore either gaining or losing energy. The photon energy is directly related to the frequency of the photon and therefore inversely to the wavelength of the radiation.

Increasing the distance between the target and the lidar instrument also increases the effect of the atmospheric absorption and scattering. In the near-IR area, the most notable atmospheric effect is the water absorption, which is has especially strong absorption bands at 1400 nm and 1900 nm, as illustrated by Figure 2. Other disturbing particles in atmosphere are aerosols, such as dust, mist, smoke, or falling snow, albeit aerosols are usually present locally or temporarily.

A lidar instrument is usually calibrated according to the air humidity, pressure and temperature in order to compensate for atmospheric scattering effects. With our instrument, a reference sample is used for this calibration; the sample of known properties is set to the same distance than the actual target, and a reference spectrum is measured before measuring the actual target. More about the calibration procedure will be detailed in Chapter 3.5. In addition to this, it is beneficial to omit the wavelength bands greatly affected by water absorption, as the signal-to-noise ratio will be poor there.

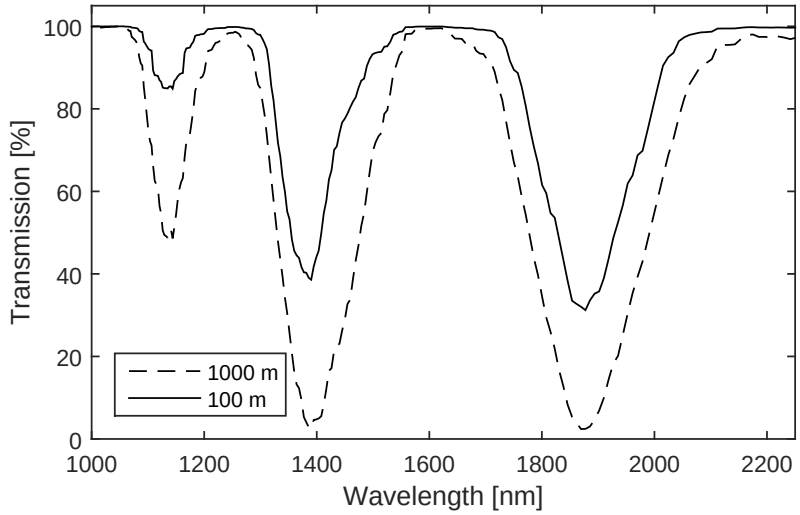


Figure 2: The absorption effect of the atmosphere at the distances of 100 m and 1000 m. Water has strong absorption bands at 1400 nm and 1900 nm. [4]

3 Experimental lidar

This work employs the hyperspectral lidar instrument that was built at MIKES in 2013 [4]. Only minor adjustments were made to the setup. This chapter introduces the lidar device, proceeding as follows: Chapter 3.1 introduces the setup as whole, offering the framework for the further detailing, while Chapter 3.2 briefly describes the used supercontinuum (SC) source and its properties and challenges. Chapter 3.3 explains the functionality of the used spectrometer, and it also features a custom-made micro-mirror spectrometer which was tested as a more affordable alternative for the more expensive commercial spectrometer. Chapters 3.2.1 and 3.4 show the only notable modifications to the physical setup, which were optimizing the SC fiber length and making the setup safe for a user's eyes by adding a dichroic mirror to physically filter out wavelengths below 1500 nm. Finally, Chapter 3.5 briefly introduces the software and data treatment.

3.1 Overall description of the setup

The used lidar instrument is featured in Figures 3 and 4. The setup is built onto an aluminum rail that is set onto a scanner, which horizontally rotates and vertically tilts the setup, enabling taking two-dimensional scan pictures of the target (three-dimensional when equipped with a rangefinder). The scanner itself is set onto a stable tripod. The output of the SC light source is delivered to the system by an optical fiber, after which the light beam was filtered out by a dichroic mirror; the transmitted shorter wavelengths are safely absorbed by a beam dump, while the rest are reflected 90^o towards the off-axis parabolic (OAP) collimating mirror. The OAP collimating mirror reflects the filtered laser beam to the target, and the backscattered light is collected by the reflective telescope. Another optical fiber transmits the collected light from the telescope to a spectrometer, where the spectral data is sent to a computer for further analysis.

The setup can be aimed at the target by either using a small off-axis aiming ocular, by using an IR camera to see the location of the beam dot at the target, or by momentarily coupling a visible light source either to the spectrometer fiber or to the SC source fiber. The setup also includes an optical rangefinder (NewkonOptik LRF MOD2) that was attached initially below the rail (as shown in Figure 3) but later moved above the dichroic mirror to bring the rangefinder closer to the optical axis of the setup.

The SC light source, spectrometer, scanner, and rangefinder were controlled by LabVIEW software from a laptop, and the same software was also used for crude spectral analysis and identification. MATLAB was used to illustrate the gained analysis results in more refined manner.

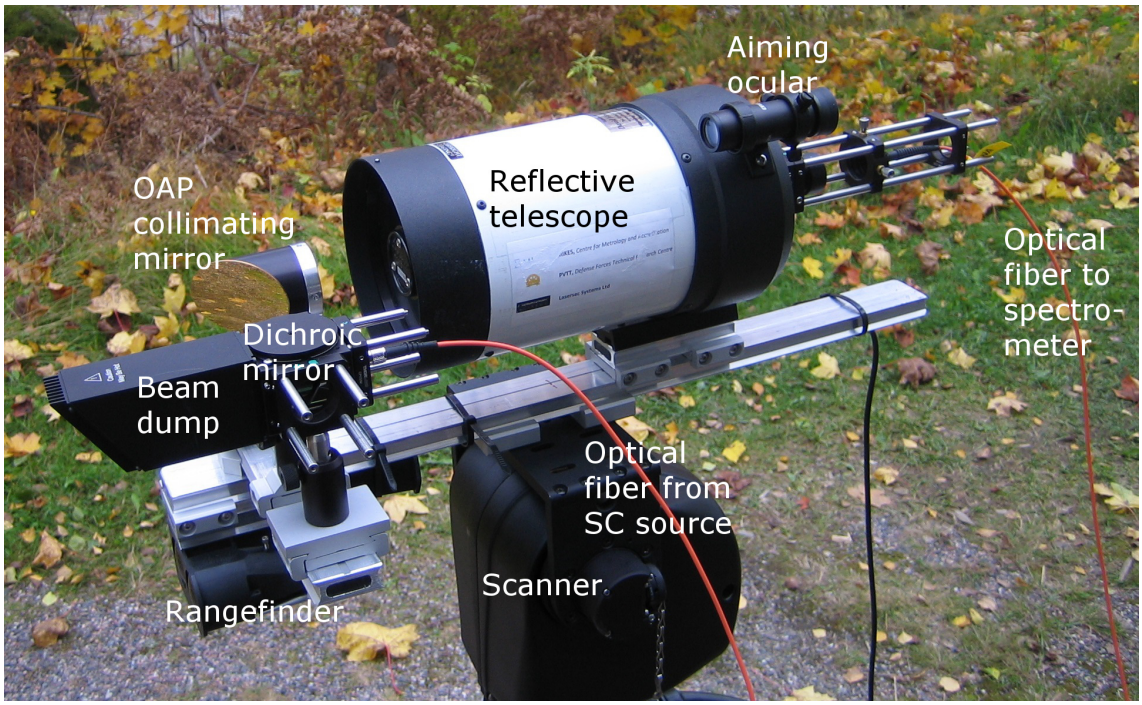


Figure 3: Photograph of the hyperspectral lidar setup used in this study. The rangefinder was later set above the dichroic mirror in order to bring it closer to the optical axis.

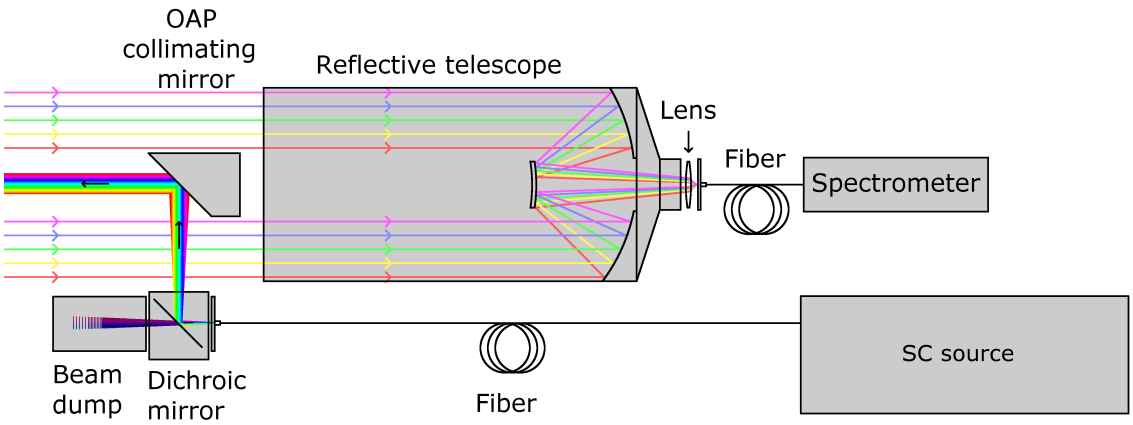


Figure 4: Schematic non-scale graph of the setup from above, with illustration of the SC and input light paths.

3.2 SC light source

A SC light source emits a board spectrum of wavelength generated by coupling a monochromatic high-intensity laser into optically nonlinear material. Usually the nonlinear material is an optical fiber, even though also a wide variety of other nonlinear media have been used [13]. Such a high-intensity SC light source has enabled the long-distance lidar measurements, as it is powerful enough to remain spatially coherent for long distances (like regular lasers) and have wide wavelength spectrum (unlike regular lasers).

In this work, the custom-made SC light source was made by coupling a monochromatic laser beam into a Corning® Infinicor® multimode graded-index optical fiber. The employed pump laser was Nufern NUQ20 fiber laser operating at 1064 nm, with pulse repetition rate of 25 kHz and peak power of approximately 10 kW. The setup was adjusted to give best possible output in terms of spectral bandwidth and intensity.

Kääriäinen's work [4] offers summarized explanation of SC generation theory as well as the development and technical details of the custom-made SC source used in this project, while a comprehensive review of SC generation theory is offered by, for example, Dudley *et al.* [13]. This thesis only explains those technical details of supercontinuum generation that are needed for further inspection.

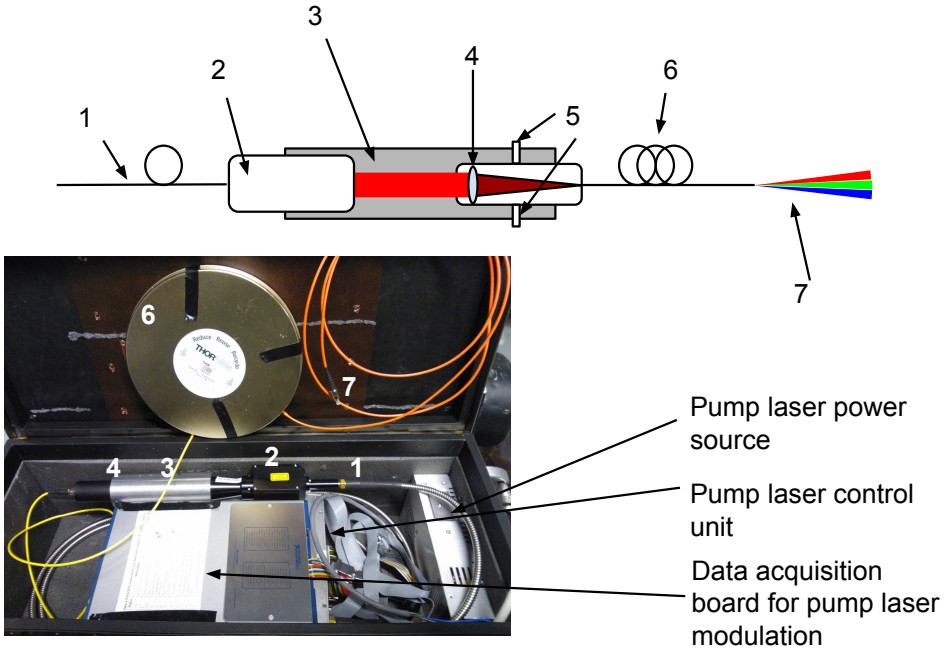


Figure 5: The SC generation setup. Pump laser beam (1) is collimated with collimator (2) and focused with a lens (4) into a 100 m long optical telecommunication fiber (6), which spreads the input laser into SC light (7). The collimator and the lens are attached to an aluminum housing (3), and the beam spot is aligned by using adjustment screws (5). [4]

3.2.1 Optimal length of the SC fiber

The first modification made to the setup was increasing the length of the optical fiber from 100 m to 240 m. The optical fiber for SC generation must be long enough to spread the output spectrum but also short enough to avoid intensity losses due to the intensity dependence of the nonlinear effects and the relatively large core size of the fiber. The original setup used a 100 m long fiber, but this choice of fiber length was not based on an experimental optimization.

In order to optimize the fiber length, the SC light output was studied by directing the original monochromatic laser input through an optical fiber that was originally 300 m long and then cut 10 m shorter between each measurement. A pair of irises was used to limit the SC output intensity in order to avoid saturation of the spectrometer. The obtained spectra are then scaled by the overall intensity of the output SC continuum in order to compare the shapes of the spectra.

The results are shown in Figure 6, and snapshots of the normalized spectra are shown in Figure 7. Figure 6 shows that beyond the fiber length of 150 m, the overall shape of the output spectrum does not change remarkably, except that the longest wavelengths slowly vanish due to the intensity losses. Increasing the fiber length also smooths the spectrum to the detriment of the output intensity. The final fiber length was chosen so that the SC spectrum would be as even as possible while the fiber length would be as short as possible. The length that fulfilled these conditions best was 240 m, and this length was used in the later setup.

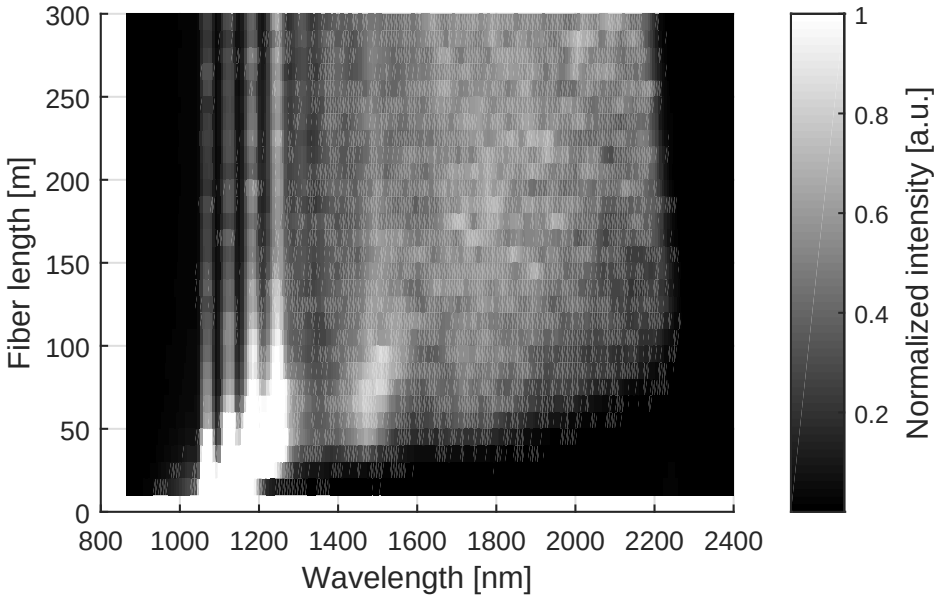


Figure 6: An intensity color map illustrating the measurements for determining the optimal optical fiber length. The spectra are scaled to have equal integrated intensity in order to compare their shapes. In addition to that, the high Raman scattering peaks are cut (even-colored white area near the lower-left corner) in order to clarify the color map.

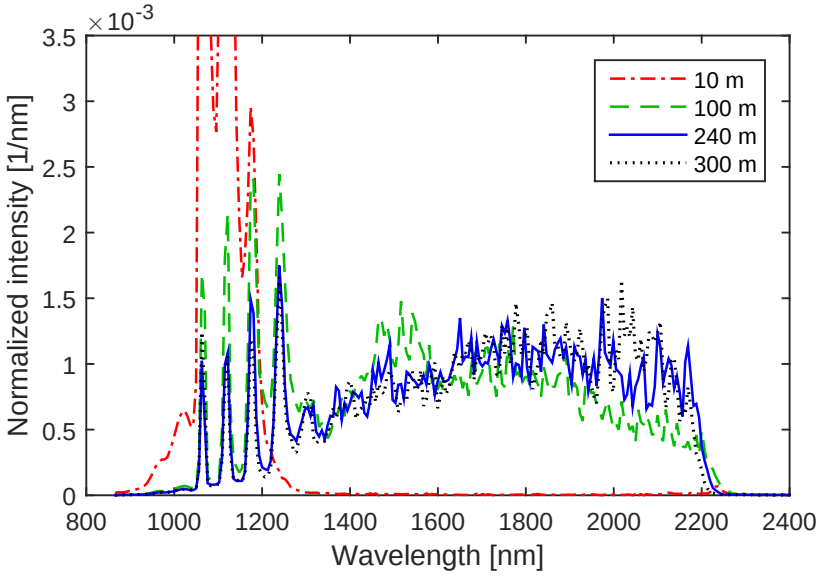


Figure 7: Snapshots of the normalized SC output spectra at different fiber lengths.

3.2.2 SC stability

The SC light source shows unstable behavior, which creates different SC spectra at different time moments even if the measurement conditions were otherwise stable. This behavior was expected, as even smallest deviations in the output of the pump laser source are amplified by the coupled non-linear optical fiber. This happens due to several intensity-related processes that may change from pulse to pulse. Drifting makes the measurements less valid to compare and causes difficulties for the identification processes. Additional measurements were performed in order to pinpoint the causes and severity of this problem.

Allan deviation was employed to study the stability of the SC output [14]. Allan deviation is typically used to determine the signal stability by dividing the data into even time windows and then calculating an average value inside these time windows, resulting in averaged data that is sparser than the original data. After this, standard deviation is calculated over the averaged data points. These averaging and deviation-calculating steps are repeated with varying time window lengths, and finally, the resulting standard deviations are plotted against the used window lengths.

The Allan deviation curve can be used to determine the ideal averaging time for measured signal. In the case of an ideal white-noised signal, the value of the Allan deviation curve decreases exponentially when the averaging time window increases, which means that greater averaging times give smoother signals. However, in real-life applications, the Allan curve starts to level off and often increase after reaching certain time window length; this is the point where the drift effects of the instrument start to dominate over random noise. Inconsistencies in the initial exponential decrease (such as plateaus) imply systematic noise elements caused by the instrument. In addition to the shape of the Allan deviation curve, the initial starting point of the curve (implying non-averaged data) is a valid noise evaluation parameter.

In this work, the stability of the lidar instrument is examined by performing a simple, 30 minutes long measurement in which a SC beam is pointed at a white reference target (featured later in Chapter 3.5) at a short distance (∼10 cm) without any optics in between. The reflected spectrum is gathered to the spectrometer from a similar distance, and the measurements were taken at 10 ms intervals. Then, a spectral angle mapper (SAM) value is calculated out of each measured spectrum. SAM is a single scalar value that expresses the difference between two spectra (in this case, the SC output spectrum and an artificial constant-value spectrum); a small value expresses small difference and a great value large difference. Therefore, SAM offers a simple method to study changes of the shape of the spectrum over time. Since SAM is also a part of one of the spectral identification methods, its mathematical theory is postponed till later in Chapter 4.3.2; for now, it is enough to know that a SAM time series indicates changes in the shape of the spectrum. Alternatively, regular standard deviation could be used for the same task, but SAM suits for our purposes better; SAM can be adjusted to average data for a channel, and therefore it ignores local noise. For now, we only employ the fact that SAM indicates the difference between the shape of the current spectrum and of the artificial even spectrum. In this phase, the SAM values were calculated using 5-fold 50 nm channel distribution at wavelengths of 1550–1800 nm and 1960–2150 nm.

The SAM time series and the Allan deviation curve made of the 30 min test measurement are shown in Figure 8. The graphs show that the standard deviation varies greatly during the first minutes of the measurement, most notably as an increase of the standard deviation during the first 100 seconds. Even if the main trend has been stabilized, there still occurs abrupt peaks in the SAM time series.

Due to the great value changes at the start of the measurements, the Allan deviation was re-calculated using only the data 18 minutes (∼1000 seconds) after the start of the measurement. Figure 9 shows the results of this calculation. There, the curve follows more a typical Allan deviation curve; initially, the Allan deviation decreases smoothly, then it increases after 10–20 seconds, and eventually the trend dissipates. The results indicate that even after keeping the setup on for some time to overcome the major changes in the start, averaging data for more than 10 seconds is unwise.

The major reasons to this instability are non-linear properties of the optical fiber: when there are even smallest deviations in the original monochromatic laser output, the deviations have a remarkable, chaotic effect on the final SC output. This is especially prominent in the hundreds of meters long optical fiber that we used, as the passing light experiences non-linear spreading for a longer time. Many low-power commercial SC lights, on the other hand, tend to have only a few meters long fibers, which decreases their instability. [13]

Another experiment was performed to determine the stability of the laser source without SC generation. A laser of lower power (approximately 3.0 W compared to the typical 20 W) was coupled into the fiber, and it was used to perform similar short-distance stability measurement than the one described earlier. The difference in this measurement is that the laser intensity is not high enough to generate SC. The results of this measurement are shown in Figure 10. We can see that excluding the

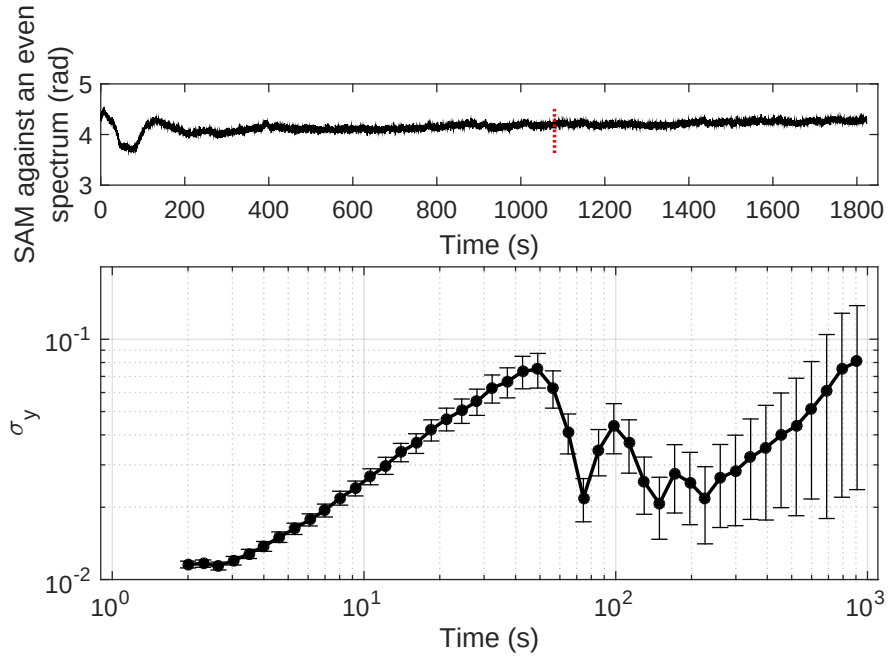


Figure 8: An example of SC drifting: the graph above features the SAM values calculated out of the spectra at each time moment, and the graph below shows the standard deviation of these values over different time frames, or the Allan deviation curve. The latter one is not very useful for these inspections, as the major value changes in the start of the measurement affect the results.

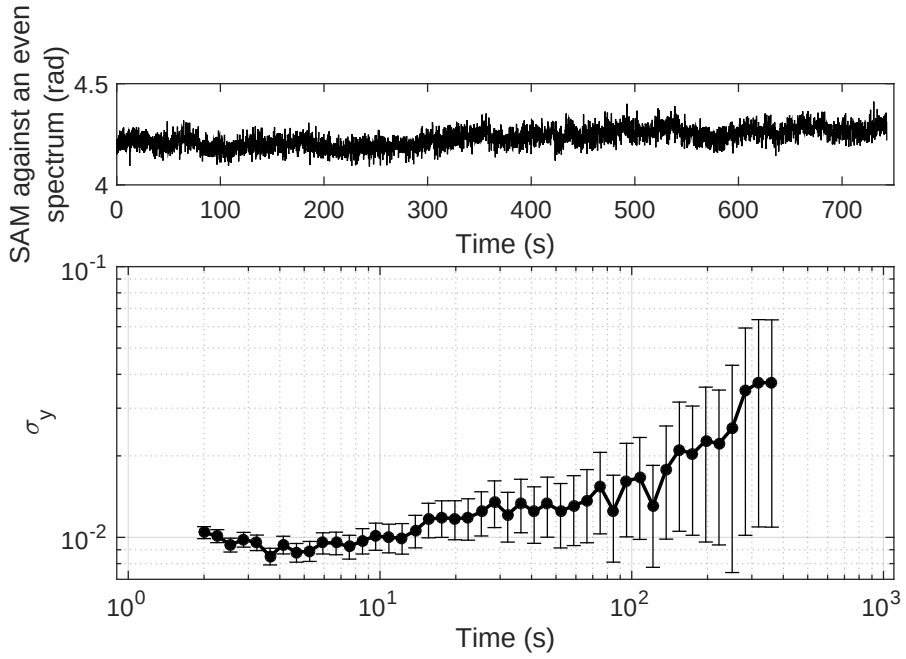


Figure 9: Results of a 30 minutes measurement after the first 18 minutes, Allan curve calculated accordingly. This Allan curve shows more expected behavior.

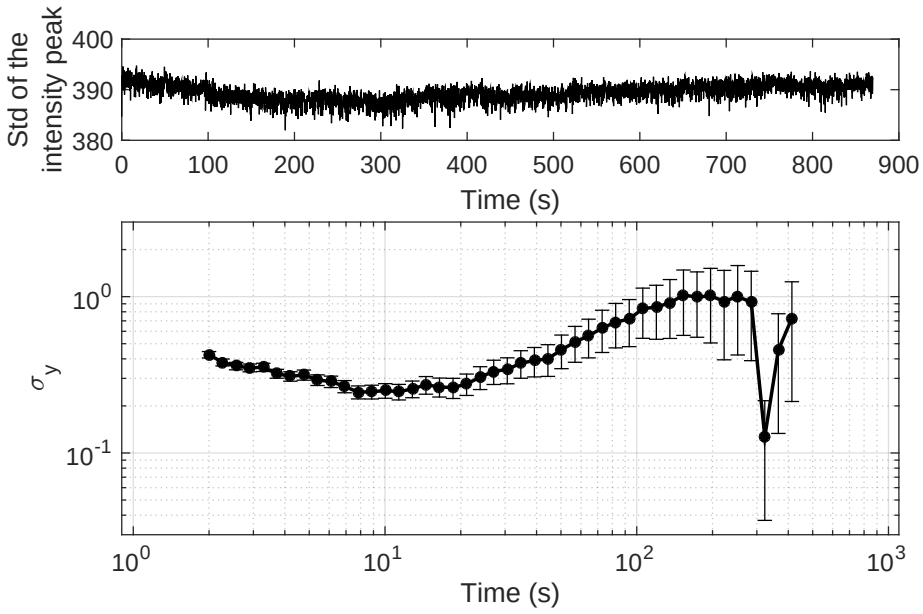


Figure 10: Measurement with 3.0 W SC light through the optical fiber, with the height of the peak as a variable.

small initial decrease, the laser output remains relatively even ovet time, confirming that a major reason for the chaotic behavior is indeed in the optical fiber.

Figure 11 contains information about how the original laser wavelength peak behaves when the laser is used with full power. The comparison between the peak height behavior (Figure 11) and the SAM value behavior of the whole spectrum (Figure 8) shows that the peak height curve expresses distortions at the same time as the SAM value curve does.

The effect of this drift was also studied by measuring spectral information from a measurement in which the SC source is turned on and off for 10 ms per either period, after keeping the SC source on for 10 minutes to avoid start-up heat gradients. The short illumination/non-illumination time period of 10 ms was chosen to test the behavior of the SC light source with extreme on/off frequency. The result of this measurement is shown in Figure 12. While this curve naturally shows periodical value changes when the SC source is flashed, but otherwise, the trend appears smoother than in the 30 minutes measurement. However, the decreasing trend of the SAM value still appears clearly.

In order to compensate for the SC drifting, the setup was temporarily upgraded with a plastic piece that was grated to diffuse the incoming light. This component, a diffuser, was moved in front of the SC output beam between each spectrum measurement, and a reference spectrum was measured in order to track the SC drifting and use it to automatically eliminate changing trend differences from the final spectrum. However, problems emerged as the diffuser did not always move to exactly the same location every time, which caused grave differences in the reference spectra. If the

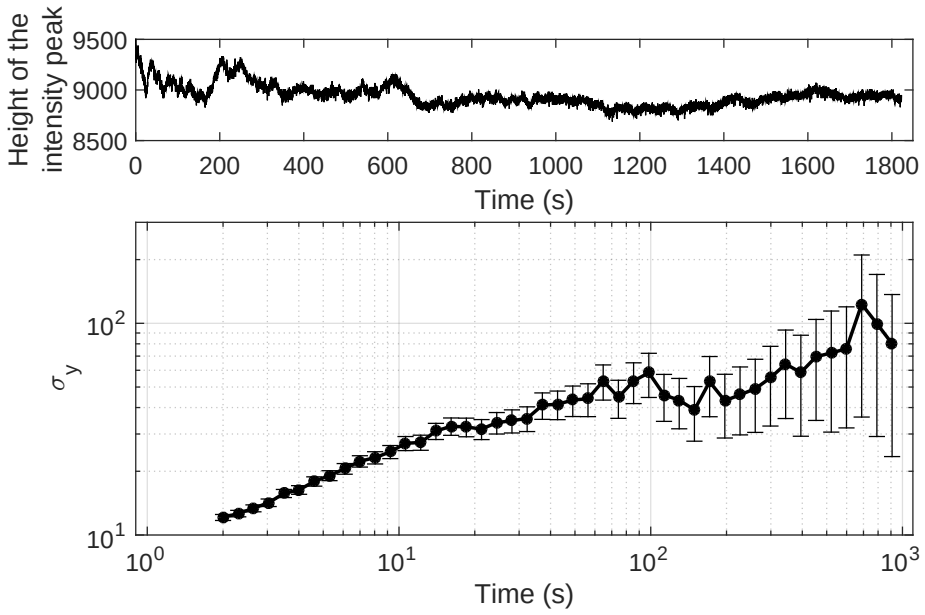


Figure 11: The height of the intensity peak at the original laser wavelength during the 30 minutes measurement. The unit of the peak height is arbitrary.

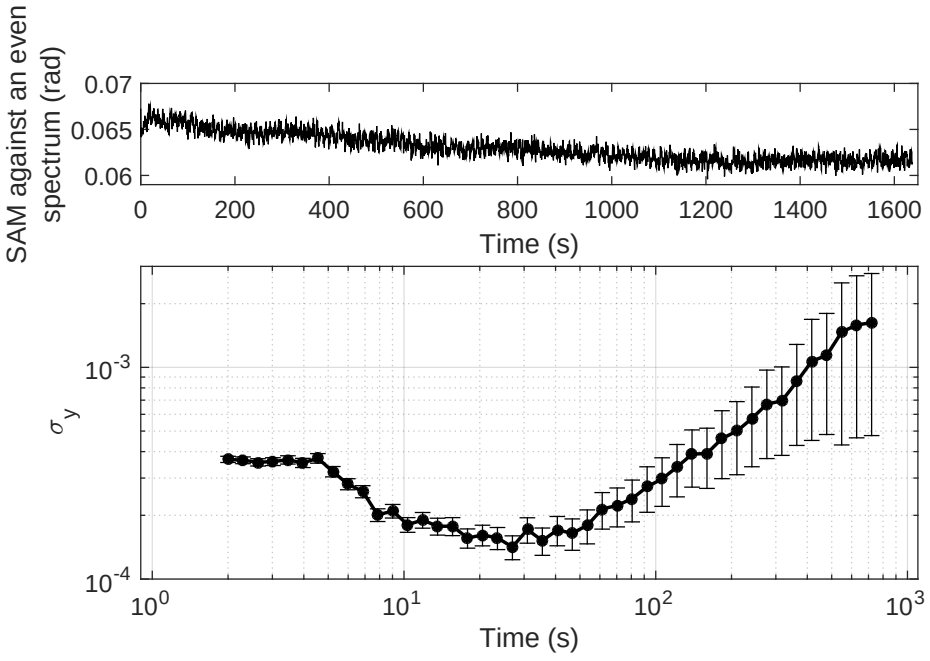


Figure 12: A measurement performed by switching the SC source on and off for 10 ms for each time.

diffuser had been perfectly diffusive, the slight spatial errors would not have caused this kind of differences in the reference. In fact, reference measurements could be improved by replacing the diffuser with one of finer grains, but the replacement did not remove the problem. The component was eventually removed from the setup, as the attempts to compensate the unstable SC behavior caused more error than the SC drifting itself. Results in Kääriäinen's work [4] also tell that the SC stability is wavelength-dependent.

Relating to the SC drift, another emerged problem in the SC stability is the effect of the movement of the optic fibers, especially the ones leading to the spectrometer. The reason for this behavior is that the changes in the fiber position also change the beam paths inside the fiber, therefore causing differences in the SC generation. When measuring single-pixel spectra, the fiber movement is not a major problem as the fibers can remain immobile, but the moving fibers damage the quality of scanned multi-pixel spectra, as the fibers must allow the movement of the scanner. During later measurements, the fibers were fastened as immobile as possible while still letting the scanner rotate.

3.3 Spectral detection

A defining characteristic of hyperspectral sensing is detection of back-scattered intensity at numerous wavelengths instead of only a few separate wavelengths or of overall intensity. For that, a hyperspectral sensing setup requires a multi-channel spectrometer with high enough wavelength resolution. This work mainly employs a spectrometer that uses Czerny-Turner design, as illustrated in Figure 13. A Czerny-Turner spectrometer uses a diffraction grating to separate the wavelengths of input light by reflecting the wavelengths into different angles and then detecting the intensities by a linear sensor array. Concave mirrors are used to collimate and focus the light beam, and an input slit is used to improve the resolution.

In this work, an Ocean Optics NIRQuest 256 grating spectrometer was used as

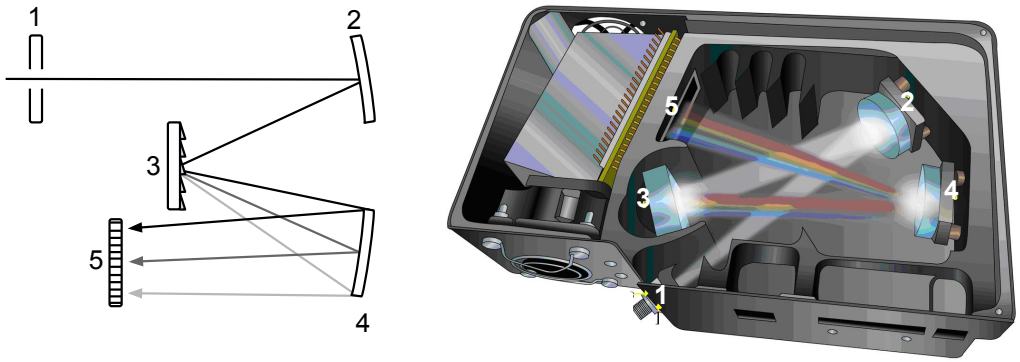


Figure 13: Czerny-Turner spectrometer design (left) and Ocean Optics NIRQuest 256 design (right). Components: (1) Input slit, (2) collimating concave mirror, (3) diffraction grating, (4) focusing concave mirror, (5) detector array. [4]

the main detector, illustrated in Figure 13. It has a 256 pixel linear sensor array of the range 900–2500 nm, and the resolution of the detector is 6.6 nm according to the manufacturer. The detector is controlled by the LabVIEW software. Since the wavelengths of interest in this study are limited to 1500–2200 nm, only 110 channels out of 256 are actually used, and even fewer when excluding the water absorption bands. The cost of the used NIRQuest spectrometer is comparable to the combined cost of the other setup components. Therefore, a custom-made grating spectrometer was built out of more affordable components, and its performance was compared to the NIRQuest spectrometer. This was done by measuring the white reference with both spectrometers.

Designed by Kraft *et al.* [15], the custom-made spectrometer differs from the NIRQuest spectrometer in that it does not have a multi-channel detector array. Instead, this design uses a steerable micro mirror, a grating, and a simple photodetector for measuring the light intensity at different wavelengths. The micro-mirror spectrometer functions as follows: after the light beam comes through the input slit into the system, the micro mirror reflects the beam onto the diffraction grating in different angles, eventually allowing the photodetector to collect only wavelengths of a limited range.

The resolution of a micro-mirror spectrometer depends on the geometric magnification and optical aberrations. There is also a tradeoff between the resolution and the throughput of the spectrometer; the micro mirror and the grating physically slice the input light into narrow bands, which limits the optical power and therefore weakens the detection. Also the small size of the micro mirror limits the passed intensity. With our micro-mirror spectrometer, the setup was optimized for 50 nm resolution, as no sharp spectral features were expected.

Figure 14 illustrates the functionality of the micro-mirror spectrometer. Figure 14 a) features the setup from above, where the white beam illustrates non-dispersed light and red illustrates dispersed light. Yellow lines feature the reflection or diffraction surfaces. The subfigure of Figure 14 a) shows the setup from side, beam colors featured similarly to the main figure. Figure 14 b) shows the ray-tracing model of the spectrometer, where the different colors imply different micro-mirror angles and *not* different wavelengths. Instead, the beam bundles emerging from the diffraction grating in different angles indicate different wavelengths.

The performance results are shown in Figure 15, in which both devices were used to measure the spectrum of the same target. The graph shows the similarity of the spectra measured by the two spectrometers, implying that the micro-mirror spectrometer works as predicted.

For now, the NIRQuest spectrometer output has higher resolution than the micro-mirror spectrometer output; a single wavelength band measured by the micro-mirror spectrometer is notably wider than one measured by a channel of the NIRQuest spectrometer ($\sim$50 nm versus 6.6 nm). Additionally, for now the micro-mirror spectrometer is vulnerable to impacts and accidental misadjustments, making it non-ideal for field measurements. The performance of the intensity detector could be improved by cooling it, as done in the NIRQuest spectrometer.

By its performance, the micro-mirror spectrometer seems promising as an al-

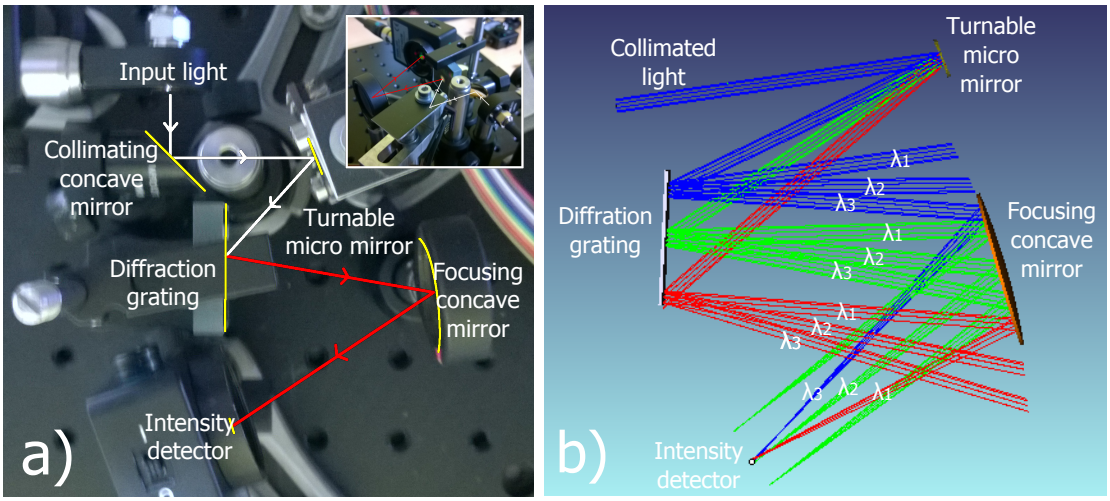


Figure 14: The micro mirror spectrometer, illustrated with a) photographs and b) a ray-tracing model.

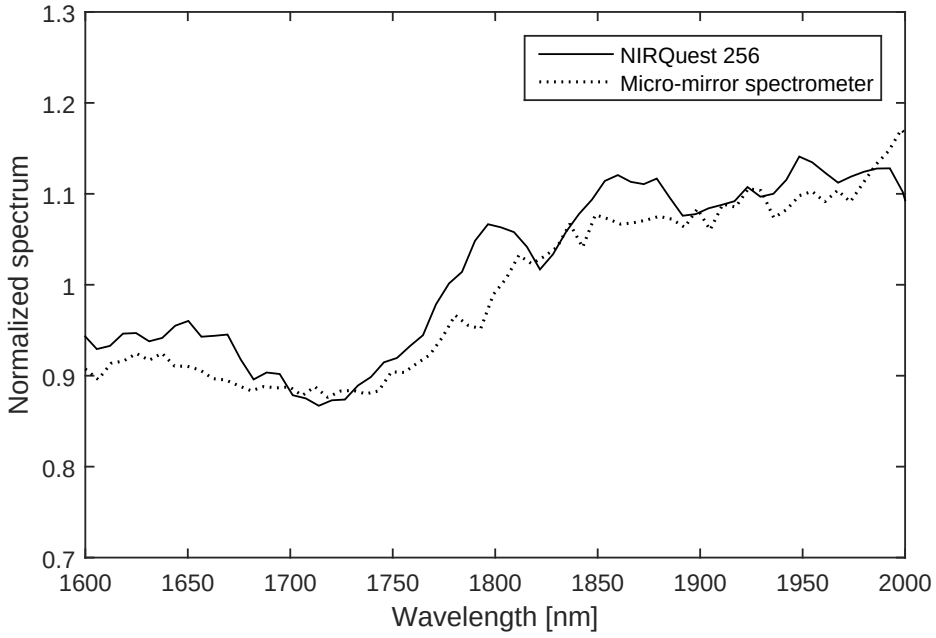


Figure 15: A performance comparison of the micro-mirror spectrometer and Ocean Optics NIRQuest 256 spectrometer.

ternative spectrometer. However, due to the limitations of the micro-mirror spectrometer, the large size of the spectrometer, the time constraints, and the fact that the lidar instrument was already programmed for the NIRQuest spectrometer, the micro-mirror spectrometer was not used in the later experiments.

3.4 Dichroic mirror

The short-wavelength peaks of the SC spectrum are most hazardous to a user's eyes, since the short-wavelength photons have greater energy than long-wavelength ones. The shortest wavelengths of the our SC spectrum are also easily transmitted into an eye due to their proximity to visible light spectrum which a human eye is specialized to detect. Meanwhile, the longer wavelengths of the spectrum are less energetic and less capable to transmit in an eye. In order to make the setup safer and enable its usage without protection gear, the dangerous shorter wavelengths are cropped out from the final output spectrum. According to the guidelines of the Radiation and Nuclear Safety Authority (Säteilyturvakeskus, STUK) [7], the filtering of wavelengths below 1400 nm is enough to enable using the instrument safely without eye protection equipment as long as direct exposure is avoided; Kääriäinen's work [4] offers more technical details concerning this wavelength filtering choice.

For filtering the shortest wavelengths, a dichroic mirror (Thorlabs, DMSP1500) was used to physically filter the high-energy low-wavelength peaks from the original SC light output. A dichroic mirror is a semi-transparent mirror that mostly reflects wavelengths above certain wavelength (in our case 1500 nm) and lets the rest transmit through. Figure 16 shows the comparison between the filtered and unfiltered SC output spectra.

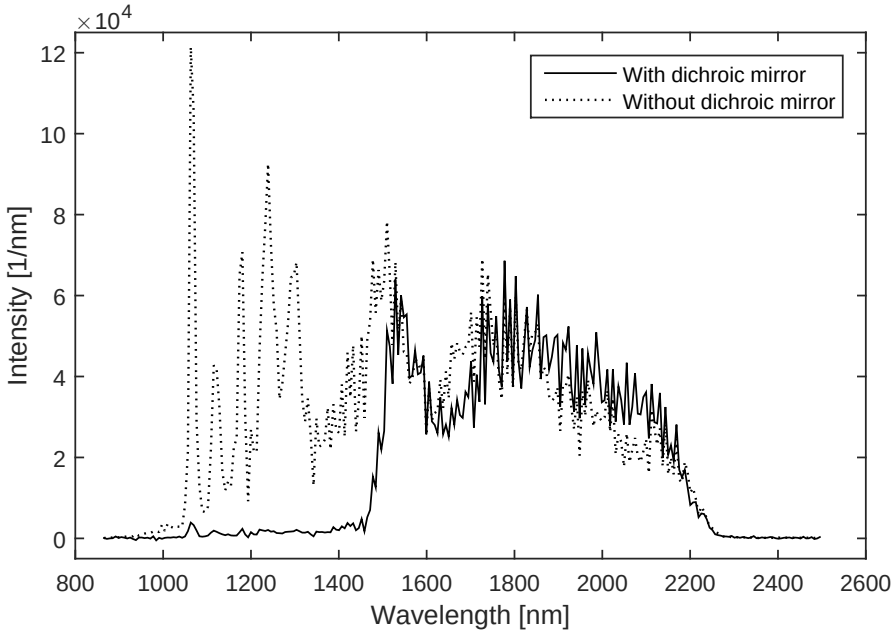


Figure 16: SC output spectrum with and without dichroic mirror filtering.

3.5 Software

Our lidar instrument was controlled by a LabVIEW software that was built as a part of Kääriäinen's work [4]. In this work, the software was implemented with spectral identification algorithms and other required data treatment functionalities. This chapter offers only a brief high-perspective overview for the software at large, including the functionalities of measuring spectra; the identification algorithms and procedures will be featured more in detail in Chapter 4.

The software offers various methods to change the settings of the spectral measurements. The pump laser power of the SC source can be adjusted, as well as the illumination time for each measurement. Typically, full pump laser power (20 W) was used to enable largest SC spread, and the illumination time was increased when the measurement distance grew. The illumination time is required to be adjusted for every measurement setting separately; too long illumination time can saturate the spectrometer, and too short illumination time periods make the signal-to-noise ratio worse. The suitable illumination time was usually set by visually investigating the spectrometer output.

Figure 17 shows the main page screenshot of the identification tab. In the current form, the software is able to treat three kinds of spectra: the latest measured single spectrum in the memory of the software, spectra saved into text files, and spectral scans saved into text files. A related functionality can also be used for saving spectra, as scan files containing also scanner position data, as detailed spot spectrum data, or as compressed spectral library data.

A single spectrum is measured as follows. Before an actual measurement is made, a calibration spectrum is measured by using a target with known reflectance properties. Our reference target is Spectralon® SRT-99-050 Reflectance Target; it is a white plastic object that is roughened and specially coated to reflect in diffused manner 92.6–98.6% of all light intensity between 250–2500 nm. The target was manufactured by Labsphere. From this onwards, the reference target will be called a white reference (WR).

The white reference is set to a distance comparable to the distance of the actual target from the setup. Then the spectrum is measured with and without SC illumination, resulting in an actual hyperspectral spectrum and a background spectrum. Subtracting the latter from the former removes the background illumination effects and results in a reflection spectrum from the white reference, or a calibration spectrum. This calibration spectrum is later used to calibrate the target spectra, therefore compensating the SC intensity variation and the atmosphere effects between the target and the lidar device. In order to remove the effect of the varying intensities of reflection spectra, the calibration spectrum is normalized so that its intensity integral over the range of 1500–2200 nm gains the value of 1 by using trapezoidal numerical integration.

Immediately after measuring the white reference spectra, the instrument is directed to the actual target, and the SC-illuminated and non-illuminated spectra are measured with the same settings to the calibration spectrum. Again, these two spectra are used to create a SC reflection spectrum out of the target, and its inte-

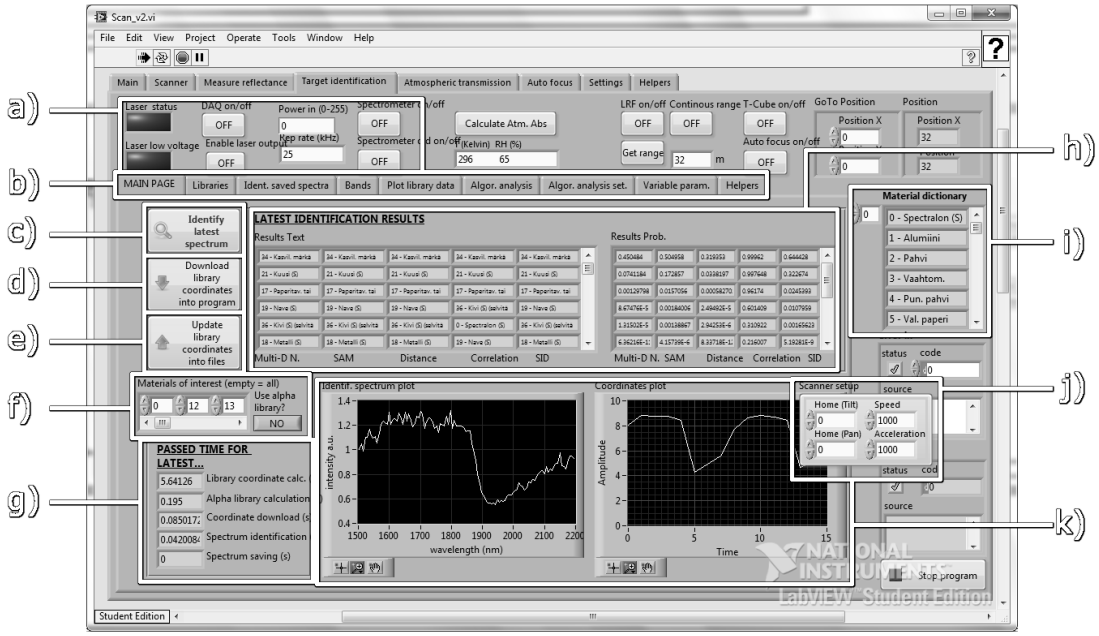


Figure 17: The main page screenshot of the identification tab. This page serves as the main control of the identification and shows the identification results. The parts of main interest are as follows: **a)** Laser and DAQ card control panel, **b)** subtabs of the identification tab, mostly containing settings for the identification algorithms, **c)** the identification button for the most recently measured spectrum, **d)** the button used to upload library spectrum coordinates from text files into the quick memory of the software, **e)** the re-calculation button for the library spectrum coordinates, used when the identification settings are changed, **f)** the material category choice as material dictionary numbers, see i), and the alpha library use choice, **g)** timers for different parts of software execution, **h)** a table of identification results, as names on left and as numbers on right, with a column for each method, **i)** the material dictionary, **j)** floating part of the scanner position and acceleration control, and **k)** images of the identifiable spectrum and its coordinate values; in this screenshot, there was used a 2-fold 50 nm band choice, which explains the vague double-shape of the original spectrum.

gral over 1500–2200 nm is normalized to 1. Finally, the measured target spectrum is divided by the calibration spectrum. In conclusion, a single final spectrum needs four different raw spectra: white reference spectra with and without SC illumination ($S(\lambda)_{WR,illum.}$, $S(\lambda)_{WR,non-illum.}$) and target spectrum with and without SC illumination ($S(\lambda)_{target,illum.}$, $S(\lambda)_{target,non-illum.}$). The final spectrum is then calculated as follows:

$$
\begin{aligned}
S(\lambda)_{final} &= \frac{S(\lambda)_{target,norm.}}{S(\lambda)_{WR,norm.}} \\
&= \frac{S(\lambda)_{target,SC}}{\int S(\lambda)_{target,SC} d\lambda} \cdot \left(\frac{S(\lambda)_{WR,SC}}{\int S(\lambda)_{WR,SC} d\lambda} \right)^{-1} \\
&= \frac{(S(\lambda)_{target,illum.} - S(\lambda)_{target,non-illum.})}{\int (S(\lambda)_{target,illum.} - S(\lambda)_{target,non-illum.}) d\lambda} \\
&\cdot \left(\frac{(S(\lambda)_{WR,illum.} - S(\lambda)_{WR,non-illum.})}{\int (S(\lambda)_{WR,illum.} - S(\lambda)_{WR,non-illum.}) d\lambda} \right)^{-1},
\end{aligned}
\tag{3}
$$

where the subtitle $final$ refers to the final spectrum, $norm.$ refers to the integral-normalized spectrum, and SC refers to the SC reflectance spectra with the background illumination effects removed. The treatment of $S(\lambda)_{target,norm.}$, $S(\lambda)_{WR,norm.}$, and $S(\lambda)_{final}$ is also illustrated in Figure 18.

The normalized spectrum $S(\lambda)_{final}$ is a relative spectrum with values that typically have an average around 1, and *it is the spectrum that is used for the actual target identification*; calibration by using the white reference SC spectrum $S(\lambda)_{WR,SC}$ enables the elimination of the atmospheric effects and therefore making different measurements comparable. It also eliminates potential intensity variations and captures only the essential information, the shape of the reflection spectrum.

Later, it was realized that the spectral normalization method is inaccurate due to the water absorption bands, especially the band at 1800–1960 nm; the integrated intensity is affected by air humidity when the water absorption bands are included. This was not corrected for this thesis, as it was noticed relatively late and as the incorrect integrated intensity values affect only the scale of the final spectrum, the multiplier being $(\int S(\lambda)_{WR,SC} d\lambda)/(\int S(\lambda)_{target,SC} d\lambda)$. The incorrect scale has only an effect on the distance algorithm but not on the other ones.

The setup was used to take two types of measurements: single-pixel measurements and scan measurements. A single-pixel measurement means illuminating the target at a single point, typically by taking several spectra one after another. These kind of measurements were usually used for obtaining test spectra and library spectra. Meanwhile, scan measurements are a series of single-pixel measurements taken at certain predetermined points so that they contain spatial coordinate data and often also distance data measured by the rangefinder.

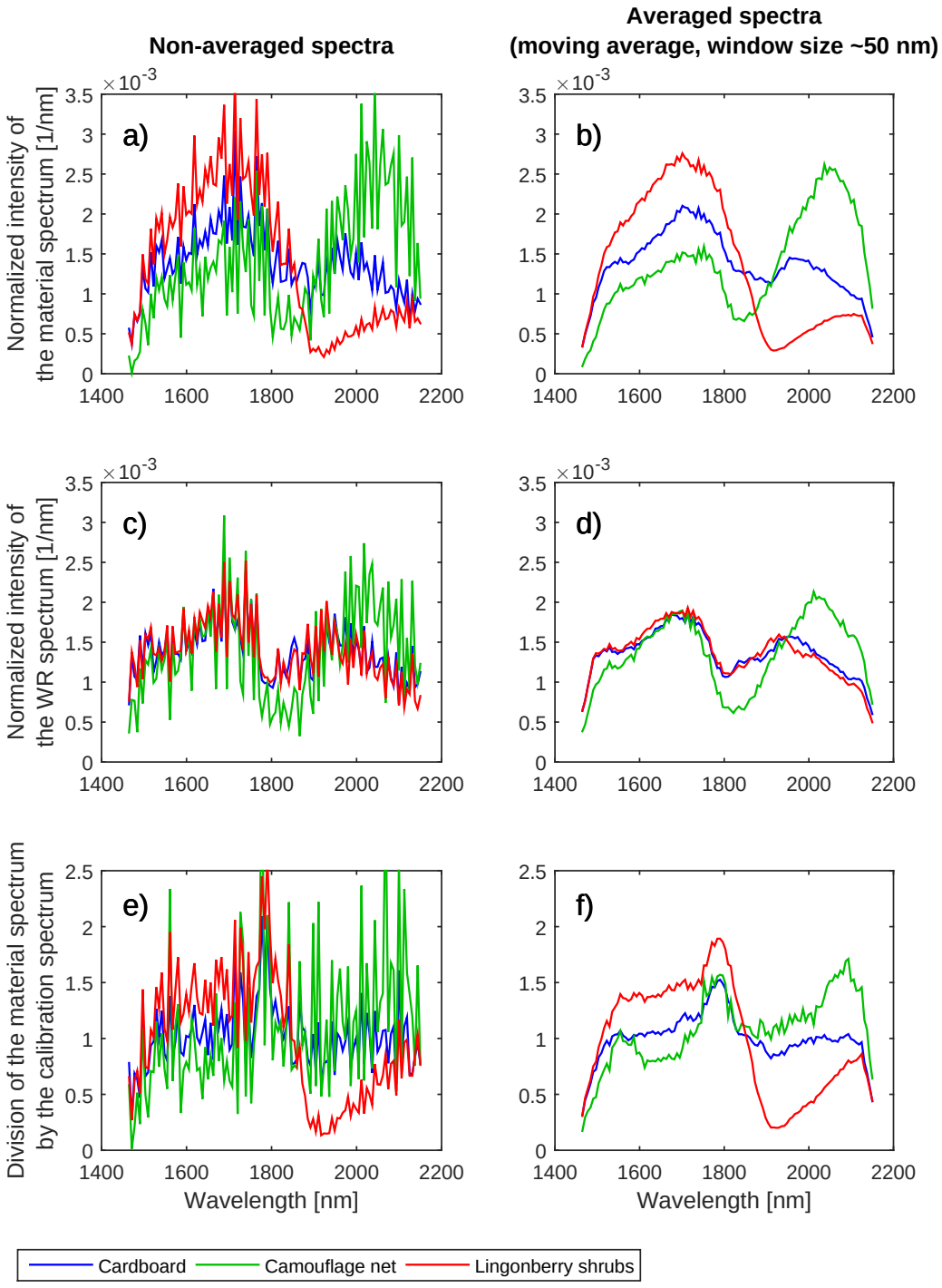


Figure 18: Illustration of spectra used in the normalization; normalized material spectrum $S(\lambda)_{target,norm.}$ in **a)** and **b)**, normalized WR spectrum $S(\lambda)_{WR,norm.}$ in **c)** and **d)**, and their division $S(\lambda)_{final}$ in **e)** and **f)**. Figures **a)**, **c)**, and **e)** show non-averaged data to illustrate the signal noise, while the data in Figures **b)**, **d)**, and **f)** are treated by calculating a moving average with time window of ∼50 nm in order to illustrate the shapes of the data.

4 Identification algorithms

The following subchapters will introduce the numerical approach of the spectral identification and its implementation into a LabVIEW software, as well as the data quality requirements.

4.1 Justification of the method choices

Several spectral identification methods are used in a variety of applications, such as geological studies, weather forecasts, and crop management. The main categories of these methods are either identifying areas of similar spectra (subpixel methods) or trying to identify single-pixel spectra by comparing them with reference spectra of known materials (spectrum matching). The subpixel methods do not require library data, therefore they suit for mapping areas with unknown materials, such as surfaces of planets. Instead, they require to have the area scanned as whole and cannot rely on single-pixel spectra. The spectrum matching requires a reference data library, but it can be used to identify single-pixel spectra. [11, 16]

Spectrum matching suits the purpose of this lidar instrument, as in military applications one is typically interested in identifying only a limited number of different known materials (such as painted surfaces), rather than in mapping areas of unknown materials. In addition to this, the final version of the instrument is expected to identify the materials while scanning, which excludes the subpixel methods as they need a whole spectral image to function. This work features five methods for spectral matching: distance, spectral angle mapper (SAM), correlation, multidimensional normal distribution (MND), and spectral information distribution (SID) methods [2, 6]. These methods were chosen due to their relevance, speed and simplicity to be written into the identification software. Since the instrument views only a small circular area at a time (diameter approximately 5 cm), we estimate that the single-pixel spectra feature mostly pure materials rather than material mixtures. Therefore, the spectral mixing is excluded from this software for now.

Naturally, these five methods are not the only spectrum matching methods. Since a limiting factor for the spectral identification is speed, iterative matching methods were excluded from the very beginning. Other potential spectral matching methods include support vectors (division of the spectral space into definite subareas according to material spectra) [17], cross correlogram method (an iterative variation of the correlation method that we use in this work) [18], and spectral feature fitting (SFF) method [11]. In addition to these, spectral identification often employs different spectral indices, and the indices will be briefly discussed in Chapter 4.4 even though they are not in the focus of this project.

4.2 Hyperspectral data as multivariable data

The identification of the spectra is performed by using multi-dimensional analysis: the wavelength-dependent intensity channels of a spectrum are divided into N bands. The intensity of each band is calculated by integrating over channels in the band, which results in N intensity values. The band widths can vary, the bands can overlap, and least interesting or worst-quality parts of the original spectrum can be omitted from the analysis. The integrated band values are used as coordinates in an N-dimensional Euclidean space, which offers a geometric base for our identification procedures. Later in this thesis, a spectrum i will be denoted as a vector $\mathbf{S_i}$, and the N-dimensional spectral coordinates made out of this spectrum are denoted as a vector $\mathbf{x_i}$.

The actual identification is performed by comparing the spectral coordinates of measured spectra ($\mathbf{x_i^M}$) with the coordinates of every spectrum in the reference data library L ($\mathbf{x_j^L}$) by using different comparison methods. Each library spectrum is associated with one or more material types T, such as "moss" or "aluminum", and possibly supercategories, such as "vegetation" or "metal". Typically, the spectral library L describes each material by a few spectra $\mathbf{S_j^L}$. Therefore, comparing the measured spectrum $\mathbf{S_i^M}$ to library spectra $\mathbf{S_j^L}$ offers information about resemblance between the spectrum and the given material type. The software can be taught to recognize new material types or to detail its knowledge of old material types by adding new library data into its library.

4.3 Employed identification methods

4.3.1 Distance method

In the distance method, the spectral measure $m_{distance,12}$ of the spectral coordinates $\mathbf{x_1}$ and $\mathbf{x_2}$ for is defined simply as a simple Euclidean distance d_{12} between the coordinates, [2]

$$m_{distance,12} = d_{12} = |\mathbf{x_1} - \mathbf{x_2}| = \sqrt{(\mathbf{x_1} - \mathbf{x_2})^T(\mathbf{x_1} - \mathbf{x_2})}. \quad (4)$$

Figure 19 illustrates this in two dimensions, which could correspond to two spectral bands. The distance information can be used for identifying the spectrum by comparing it to the spectral coordinates of known measurements: the closer the identifiable spectral coordinates are to the library coordinates, the more similar the spectra are. The measured distance d_{12} is compared to a one-dimensional normal distribution, which gives an estimate for the resemblance between these spectra. The mean μ_N of the normal distribution is 0, and the variance σ_N^2 is evaluated semi-manually by scaling the variance of the library spectrum with an experimental scaling parameter. This method assumes that the measured spectrum exhibits only minor variance that is relatively equal in all wavelengths. However, the variance of the library spectrum can be used for evaluating the resemblance between two spectra at given distance d_{12}.

The method also assumes that the measured and library spectra were taken in similar illumination conditions (such as exposure to similar intensity), or scaled so

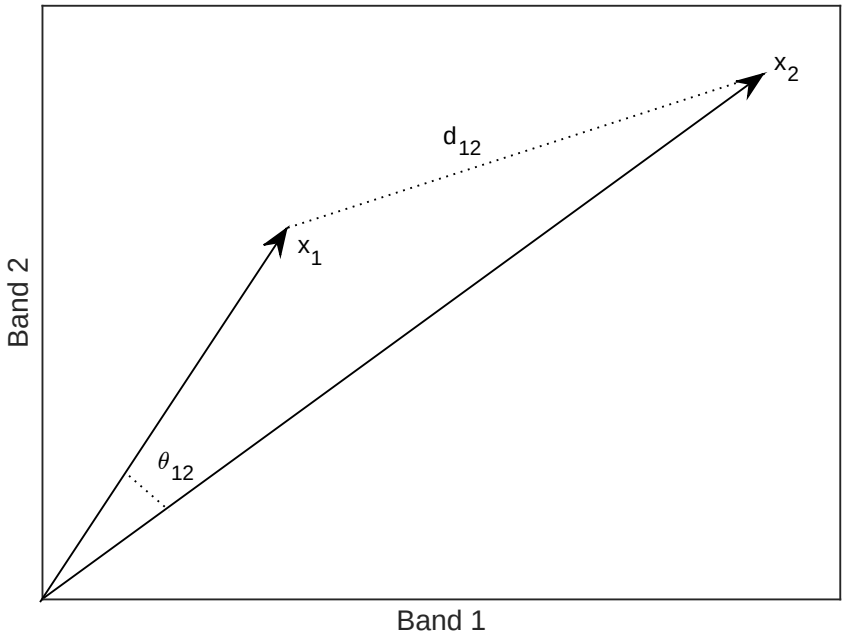


Figure 19: Visualization of the distance d_{12} in the distance method and the angle θ_{12} in the angular method, with N = 2.

that the illumination condition differences are eliminated. In this work, this is done by integral-normalizing in the 1500–2200 nm band, therefore eliminating the effect of varying intensities and concentrating on the varying shapes of the spectra. However, later this choice was found ineffective due to the effects of the water absorption bands, as is later stated in the results.

4.3.2 Spectral angle mapper (SAM) method

In the distance method the spectral coordinates depend on the intensity of the employed light source: the brighter the illumination is, the greater is the intensity of the measured spectrum and the greater coordinate values are calculated. However, the *direction* of the $\mathbf{x_i}$ coordinates in comparison to the N-space origin is not affected by the original intensity, if the intensity is assumed to affect the measured spectrum linearly. This angle can be thought of as a generalized angle,

$$m_{SAM,12} = \theta_{12} = cos^{-1}\left(\frac{\mathbf{x_1^T x_2}}{\sqrt{(\mathbf{x_1^T x_1})^T(\mathbf{x_2^T x_2})}}\right). \tag{5}$$

An example of angle θ_{12} is featured in Figure 19. The angle θ_{12} is the basis of the spectral angle mapper (SAM) method developed by J. W. Boardman [12], and the method is widely used in spectral identification due to its simplicity and independence from the light intensity [11]. Like the distance method, the SAM method assumes the variances of the band insignificant, and usually, the method uses averaged

library samples [19]. However, this kind of samples do not express the variability of the spectra. However, in 2014 Zhang and Li attempted enabling variability in SAM-based spectral identification by (1) using spectral derivatives and (2) calculating θ_{12} for several library spectra of the same material and choosing the one with smallest value [19]. Our approach also uses several spectra, though using them to form a spectral probability cloud rather than treating them as single spectra.

In this work, the variance data is used for evaluating the resemblance between two spectra at given angle θ_{12}. Similarly to the distance method, the similarity between two spectra is determined by using a normal distribution scaled with the average of the coordinate variances of the library spectrum.

4.3.3 Correlation method

The correlation method compares two coordinate sets as if they composed two-dimensional linear data, so that one set corresponds to the values on X axis and another one values on Y axis. The method calculates a simple Pearson correlation test between the spectra

$$\begin{aligned} m_{correlation,12} &= 1 - r = 1 - \frac{cov(x_1, x_2)}{\sigma_{x1}\sigma_{x2}} \\ &= 1 - \frac{\sum_{k=1}^{n}(x_{k,1} - \overline{x}_1)(x_{k,2} - \overline{x}_2)}{\sqrt{\sum_{k=1}^{n}(x_{k,1} - \overline{x}_1)^2}\sqrt{\sum_{k=1}^{n}(x_{k,2} - \overline{x}_2)^2}}, \end{aligned} \tag{6}$$

where r is Pearson product-moment correlation coefficient and $\overline{x}_1$ and $\overline{x}_2$ are mean values of the spectral coordinates. In the original Pearson correlation test, the data sets gain value $r = 1$ if the sets have perfect positive correlation, or they can fit onto an ascending straight line flawlessly, and value $r = -1$ if they have perfect negative correlation, or they can fit onto a descending line. Therefore, trying to create a line out of coordinates of x_1 and x_2 could offer a simple method to evaluate their similarity, as illustrated in Figure 20. In the case of negative r values, r reduces to 0, as obviously a descending line would not imply similar spectra at all.

Der Meer and Bakker [18] also introduced the cross correlogram spectral matching method, which can be seen as an iterative version of the featured simple correlation method. In the cross correlogram method, the measured spectrum and the library spectrum are gradually shifted against each other, and the correlation values are plotted as a function of these step shifts. In a matching case, the resulting curve should have the largest value when no shifts are made and gradually decrease when the shift increases. The matching value would then depend on the highest correlation value and the overall form of the correlation curve. However, the cross-correlation method was excluded from the scope of this work, as it would be more time-consuming than the other methods due to its iterative nature. In addition, using spectral shifts would not be intuitive for our measurement setup.

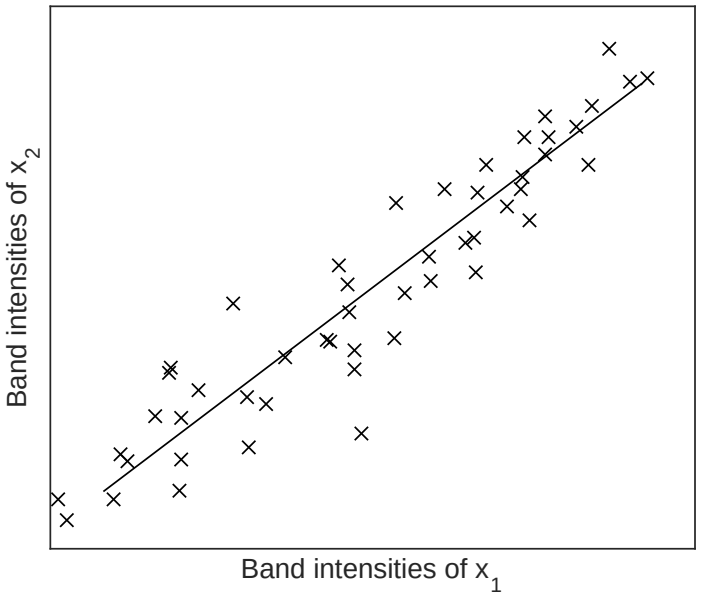


Figure 20: An example of applying the correlation method for differentiating spectra $\mathbf{x_1}$ and $\mathbf{x_2}$. Each cross stands for a pair of corresponding bands from the two compared spectra.

4.3.4 Multidimensional normal distribution (MND) method

Actual hyperspectral measurements have error factors, such as limitations in the device. In addition, spectra of similar material samples may vary. Therefore, it can be reasonable to model different kinds of materials as N-dimensional *material probability densities* $p(\mathbf{x_i})$ rather than as single point coordinates. Single data points can be estimated as N-dimensional ellipsoid-shaped normal distributions. Therefore, similarity between two spectra can be estimated by forming a density function out of one spectrum and calculating the density value at the location of the other one; typically, a library spectrum forms a density function and the measured spectrum is compared to it. For example, spectral coordinates $p(\mathbf{x_2})$ form a data cloud, and the density value of spectral coordinates $p(\mathbf{x_1})$ can be calculated as

$$\begin{aligned} p_{MND,12} =& p_{MND,\mathbf{x_2}}(\mathbf{x_1}) = f_{norm,ND}(\mathbf{x}_1, \mathbf{x}_2, (\sigma_2 \cdot VP)^2) \\ =& \frac{1}{(2\pi)^{N/2}} \frac{1}{|\mathbf{\Sigma_2}|^{1/2}} exp\{-\frac{1}{2}[\mathbf{x_1} - \mathbf{x_2}]^T \mathbf{\Sigma_2}^{-1} [\mathbf{x_1} - \mathbf{x_2}]\}, \end{aligned} \tag{7}$$

where $f_{norm,ND}(x, \mu, \sigma^2)$ is an N-dimensional normal distribution density function, $\mathbf{\Sigma_2} = (\sigma_{x2} \cdot VP)^2 \mathbf{I}$ stands for variance matrix for the scaled variation of $\mathbf{x_2}$ in each dimension, while $|\mathbf{\Sigma_2}|$ is the determinant of $\mathbf{\Sigma_2}$. The integrated value of this function over all the space is 1 [2]. Figure 21 shows a two-dimensional example of the $p(\mathbf{x_2})$ probability cloud and how $\mathbf{x_1}$ is compared to it.

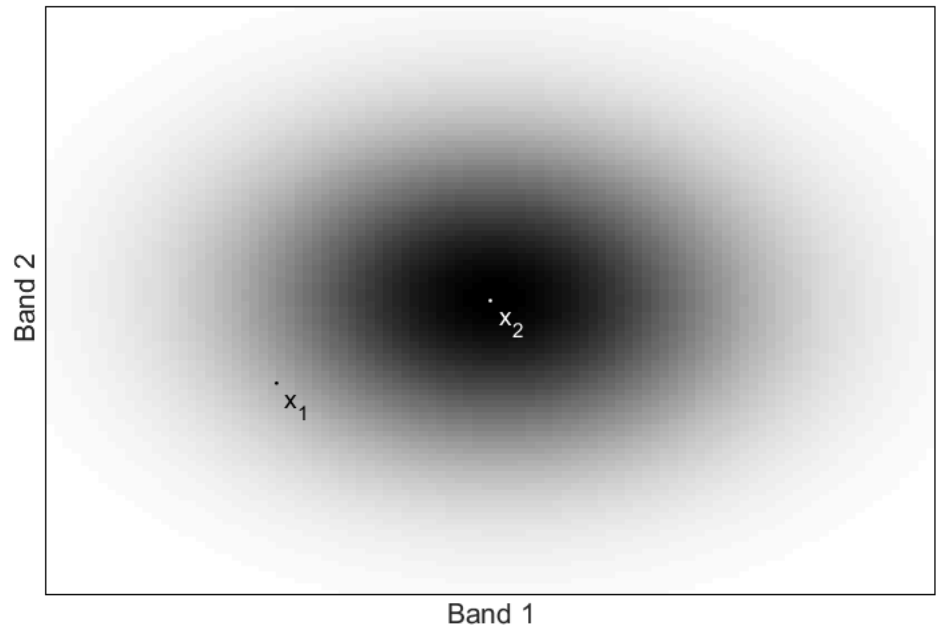


Figure 21: A visualization of two-dimensional normal distribution probability cloud of $\mathbf{x_2}$. The probability density value p_{12} is determined by the location of $\mathbf{x_1}$ inside the probability density cloud formed by $\mathbf{x_2}$. In this case, the Band 2 of spectrum 2 had less variation than Band 1, affecting the shape of the cloud.

4.3.5 Spectral information divergence (SID) method

Spectral information divergence method is a fairly recent spectral identification method, proposed by Chang in 2000 [6]. It is based on a new hyperspectral measure named Spectral Information Measure (SIM), which describes the variability of the spectral coordinates rather than spatial location. Chang suggested in his article that SID would be a more effective spectral identification tool than the widely used SAM method, at least according to his calculations that tested the differentiation capability of the methods. Later in this thesis, we introduce and use the differentiation tests employed by him, or rather, altered versions of those tests.

SIM is based on the expression

$$H(\mathbf{x_i}) = -\sum_{l=1}^{L} p_l log p_l, \tag{8}$$

where p_l is the probability measure of a channel l, $p_l = x_l / \Sigma x_j$. The calculated $H(\mathbf{x})$ contains spectral information of the spectrum, namely the sum of its mean, variance, and central moments of 3...∞ order. This, in turn, can be used to determine the relative entropy of a spectrum $\mathbf{x_1}$ with respect to $\mathbf{x_2}$:

$$D(\mathbf{x_2} \parallel \mathbf{x_1}) = \sum_{l=1}^{L} p_l log \left(\frac{p_l}{q_l} \right), \tag{9}$$

where p refers to the probability measure of $\mathbf{x_1}$ and q of $\mathbf{x_2}$. The final result, SID, is a spectral measure that is defined as

$$m_{SID,12} = SID(\mathbf{x_1}, \mathbf{x_2}) = D(\mathbf{x_2} \parallel \mathbf{x_1}) + D(\mathbf{x_1} \parallel \mathbf{x_2}). \quad (10)$$

4.4 Spectral indices

In addition to modelling the spectra using multidimensional coordinates, it is possible to highlight the relationships between certain wavelengths by calculating a variety of indices out of the data. Some of the best known of these indices are different vegetation indices that can be used for differing organic material from inorganic, one of such indices being Normalized Difference Vegetation Index (NDVI) developed by Rouse *et al.* [20]. A single spectral measurement can be used for calculating several different indices at once, and the process can be automated by ease. However, our possibilities to use readily available indices are limited, as they are usually featured at wavelengths shorter than 1000 nm. [21]

This thesis does not focus on finding or employing spectral indices, but nevertheless we were able to identify wavelengths that could be used for such purpose. For example, vegetation shows a steep slope in the 1900–2100 nm wavelength band, decreasing towards longer wavelengths, and wet vegetation has a steeper slope than the dry one.

4.5 Application of the identification methods

When comparing two spectra, all methods except the MND method produce a spectral measure $m_{method,ji}$, which indicates the similarity between two spectra. The smaller $m_{method,ji}$ is, the more similar the spectra are. This measure is then compared to a zero-centered one-dimensional normal distribution, which has a variance determined according to the variance of one of the spectra (in our case, the library spectrum $\mathbf{x}_j^L$) and a semi-manual variance scaling parameter (later variance parameter, VP). The resulting density value p_{ij} implies alternatively scaled similarity between spectra, now highlighting the differences between almost similar-looking spectra and adjusting the result according to the variance of the library spectrum. The MND method gives directly the density value p_{ij} instead of a spectral measure.

The density values between the measured spectrum and the spectra of the same material type T_k are used to form a summed data cloud. When the measured spectrum $\mathbf{x}_i^M$ is compared to all library spectra $\mathbf{x}_j^L$ of the same material type ($j \in T_k$), the density values are summed and then scaled with the (estimated) maximum density of the summed density function, $max(\sum p_{ij})$, $j \in T_k$. This results in an identification certainty value to the material type T_k, $P_{T_k}(\mathbf{x}_i)$. The parameter $P_{T_k}(\mathbf{x}_i)$ ranges between 0 and 1 so that 0 implies absolutely no similarity between the spectra and 1 implies a perfect match. The primary guess of the identification is the material type that gives the highest $P_{T_k}(\mathbf{x}_i)$ value. This can be expressed mathematically as follows

$$
\begin{aligned}
T_{\text{best guess}}(\mathbf{x}_i) =& T_k, \quad T_k \in L, \\
& P_{T_k}(\mathbf{x}_i) \geq P_{T_l}(\mathbf{x}_i), \\
& T_l \in L, \\
P_{T_l}(\mathbf{x}_i) =& \frac{\sum p_{ij}}{max_{estimated}(\sum p_{ij})}, \quad j \in T_l, \\
p_{ij} =& f_{norm,1D}(m_{method,ji}, \mu, \sigma^2) \\
\mu =& 0, \\
\sigma^2 =& (\overline{\sigma}_j \cdot VP)^2
\end{aligned}
\tag{11}
$$

where $f_{norm,1D}(m_{method,ji}, \mu, \sigma^2)$ is one-dimensional normal distribution density function and $\overline{\sigma}_j$ is the average standard deviation over the bands of library spectrum $\mathbf{x}_j^L$. As seen, despite the name "variance parameter", VP is actually used for scaling the standard deviation rather than the variance. The parameter was indeed originally used for scaling variances, and therefore the name stuck.

4.6 Probability clouds

After identifying the probability densities between the test spectrum $\mathbf{S_i^M}$ and each library spectrum $\mathbf{S_j^L}$ of certain material type T ($j \in T_k \in L$), these densities are summed together and normalized with a scaling parameter that corresponds to the maximum density of the summed probability cloud. This treatment was expressed by Equation 11. This section describes how the estimated maximum density is calculated, and also the problems with the summed probability clouds are discussed.

The density clouds can be visualized as seeing several spectra of a certain material type as an N-dimensional probability cloud. The clouds are calculated by using the chosen library spectra and an identification method that is used to determine the probability density function of each single spectrum. The shape of the cloud also depends on the settings used for the identification method, such as the band division and variance scaling parameters. Using probability clouds is useful, especially if a single material type can have multiple types of spectra; unlike the coordinates of a single spectrum, the shape of the cloud is not limited to only ellipsoidal shapes, and it can form even probability islands. In fact, the measured spectra are compared to material probability clouds that are made of library reference spectra.

A probability density cloud for a material type T_k is formed by summing together the probability density functions of all spectra of the given material and then scaling the probability values so that the greatest value of the cloud is 1, by dividing the cloud by the maximum value of the density cloud. In reality, however, the maximum value is determined by going through each spectral coordinate in the data cloud and scaling the cloud according to the highest density value. This may cause probability densities higher than 1 in certain parts of the probability cloud, if two sets of spectral coordinates are close to each other. However, going through spectral coordinates is likely to be the easiest and fastest way to determine the maximum density value, especially when N increases.

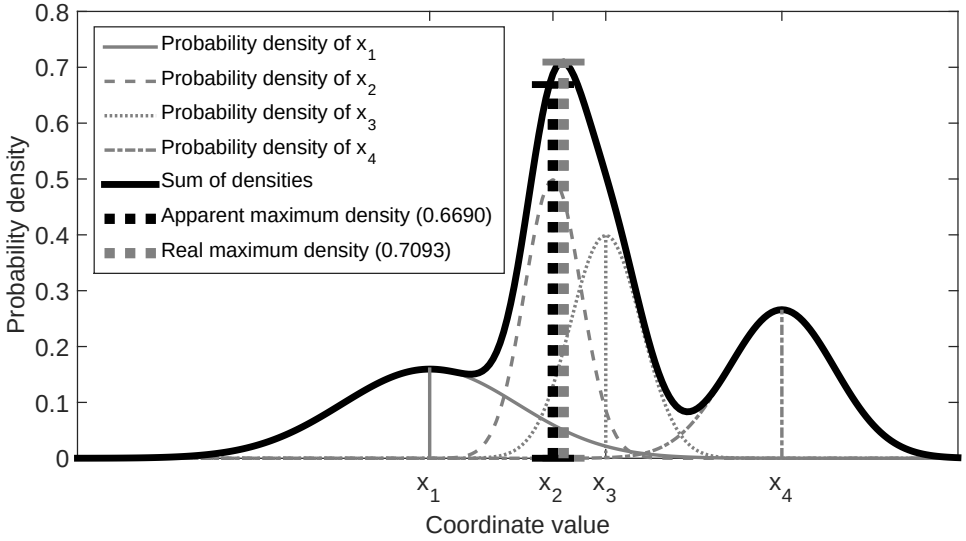


Figure 22: An example of forming a probability cloud from four one-dimensional ($N = 1$) spectral coordinates $\mathbf{x_i}$ and scaling it.

Figure 22 shows a one-dimensional example of the scaling. First, the sum probability density function is created by using the density functions of each spectral coordinate $\mathbf{x_i}$, and after that the value of the summed density function is verified at each employed spectral coordinate. In this example, the summed probability density function has greatest value at the coordinate $\mathbf{x_2}$, and probability cloud is scaled by using maximum density value, which in this case is the summation probability value at $\mathbf{x_2}$, 0.6690. However, the real maximum value, 0.7093, is located between coordinates $\mathbf{x_2}$ and $\mathbf{x_3}$.

Problems with probability cloud forming emerge if the cloud uses spectra that are located very sparsely or unevenly; for example, if there are a few spectra very close to each other and other few further away, the dense bundle of spectra gains more weight than the other parts of the probability cloud. This is not a problem if all used spectra are similar (close to each other in the N-dimensional space compared to their scaled variances), but if there is larger variability among the given material, uneven spectra highlight the correctness of densely located spectra and underestimate sparser yet equally correct spectra. This problem can be minimized by using a comprehensive collection of spectra and limiting the material categories so that they contain mostly similar spectra (such as making several "plastic" categories rather than forcing them all into the same supercategory).

4.7 Band division choice

The coordinates $\mathbf{x}$ depend on the band division of the spectra. A successful band division allows the software to use most crucial wavelength bands and ignore least informative wavelengths in order to improve the identification performance and to reduce calculation time. The band division is also dense enough to allow accurate

identification but sparse enough to avoid too many coordinate dimensions and, therefore, too complicated analysis and risking paying attention to non-differentiating and hence unimportant bands. Excluding unimportant bands could highlight the differences between spectra.

Due to the time limitations, we did not focus on honing the band division choices, but instead, we aimed to draw some general guidelines. We first settled down with wavelength bands of 50 nm, as averaging data with a window of this length clearly smooths the data well enough (as seen in Figure 18). Also, we excluded bandwidths that were strongly disturbed by water absorption, most importantly the band 1800–1960 nm [22]. We later excluded also channels below 1550 nm and above 2150 nm, as these wavelengths were also disturbed by the atmospheric effects (see Figure 2). The bands were set evenly onto the useful channels at either every 50 nm, 25 nm, 12.5 nm, or 10 nm, of which in three latter configurations the bands overlap, forming two, four, and five-fold band configurations.

However, the even band divisions may also include wavelength channels that express little difference between materials and belittle channels with major variability. If the importance of different channels is known, the significance of important bands could be highlighted by using multiplying constants and unimportant bands could be omitted. However, this was not in the focus of this thesis.

A method to determine the most crucial data bands could be as follows: the library data could be averaged by each known material type, then coordinates could be calculated out of the average spectra while using dense, even band division, and then variances of each dimension could be calculated over all the spectra. The high variances of these coordinates would imply that the averaged material type spectra differ greatly at these variances, making these bands important for identification. Similarly, low variances would imply that the material types have little differences at given wavelengths. This functionality was implemented into the software, but the more refined band division analysis was not performed within the scope of this work.

As is later pointed out in this thesis, the need of band split should be evaluated. This can be done not only by testing band splitting methods experimentally, but also by studying known spectra visually and trying to determine the most important differentiators between different materials, or by determining the most crucial wavelength bands.

4.8 Non-alpha library and alpha library

The software uses the library spectra in two ways; either it uses all spectra separately as building material for probability clouds, or it averages them first into fewer spectra and then calculates the clouds. In this work, the spectral library used for the former method is called the non-alpha library and the one used for latter is called the alpha library. The alpha library uses the non-alpha library as its source material. Later in this work, we compare the usage of these two libraries.

In the non-alpha library approach, the library spectra are made of single spectral measurements, usually without averaging. Since long enough spectral time series

were not measured, the actual variability of the spectra is often unknown or poorly known. Therefore, the library spectra use artificial constant variances of 1, equalling the typical magnitude of a spectral channel; the actual magnitudes are adjusted by multiplying them with variance parameters, which also affects the shapes of the probability clouds. The abundance of the library spectra is used to compensate for the uncertainty of the actual variances of the spectra and non-averaged spectra in general; the library spectra form a probability cloud, and their slightly altering locations indicate differences in the variability of the spectral variances. Measuring longer time series for each library spectrum and then averaging it would have been a valid action to take, but its importance was not realized until later. This is another topic to improve in the further development of the lidar instrument.

Performing identification with the non-alpha library may suffer from long calculation time, because the identifiable spectrum is compared to several library spectra one by one. While adding spectra into the library makes the identification more detailed and versatile, it also increases the number of spectra to be compared, and therefore it also increases the calculation time. An alternative way to model all the library spectra is to average several spectra measured from a single target – or, forming an alpha library. The alpha library is formed by averaging the spectra that have been taken from the same target, using the variance of these spectra is used as the variance of the final spectrum.

The usage of the alpha library allows decreasing the number of library spectra. Technically, this also reduces the spectral identification from the "spectrum compared to the probability cloud" method to the "spectrum compared to a single low-noise library spectrum" method that is a more typical approach to the most algorithms mentioned earlier. However, the alpha library often features a few spectra of each material type, as the averaging is performed for the spectra taken from the same target, and there may be several same-material targets.

We expect that the non-alpha library suits the identification software when it is taught a new material; the alpha library suits when identifying a well-known material. It could also be possible to mix these libraries, albeit that was not tried in this work.

4.9 Minimizing the calculation time

A short calculation time is essential for applying the identification software for practical uses. Specifically, the calculations are required to be fast enough to function as a part of spectral scanning, or in other words, to perform identification faster than the device performs scanning. The calculation time is affected mainly by computation capacity of the hardware, execution time of a single two-spectrum comparison, spectral band division, library data access, and most importantly, number of spectra in the identification library. In addition, the time required to measure and average a clear spectrum affects the identification time.

4.9.1 Optimizing band division

While splitting the spectral data into even higher number of bands (or, increasing N) makes the spectral analysis more refined, too many dimensions make the data analysis more complicated. Also, slicing spectra into too narrow bands makes the bands more vulnerable to noise and other error terms. These issues were detailed in Chapter 4.7, which also features our band division choices. However, the band division does not have a great effect on the calculation time; theoretically, increasing N increases the time to calculate the spectral coordinates, but this increase is insignificant compared to the time required for opening, reading, writing, and closing library spectrum files. Band division does not significantly affect the actual two-spectra comparison time, as the coordinates are treated as vectors and no dimension-dependent iterations happen.

4.9.2 Reference data access and number of library spectra

The calculation time of the identification process can be reduced by minimizing the time used to handle the library data. Before the identification processes, the band division choices are made, and all the spectra are converted into coordinates and their variances and saved into separate text files with minimized sizes. The coordinate data is then downloaded into the identification software as vectors. Naturally, the number of library files affects directly the calculation time.

The coordinate calculation and data conversion are the operations that take most time out of the calculation processes; with 189 library spectra, it takes approximately 5–7 seconds to perform coordinate calculation for non-alpha library (less than 0.5 seconds for alpha library) and approximately 0.1–0.2 seconds to perform data download. However, it is irrelevant to attempt to optimize the coordinate calculation and download time, as the operation is needed to be performed only once when the lidar instrument is switched on – or, in rare cases, when the band division and/or spectral identification parameters are altered.

4.10 Software flexibility

Each known spectral data point is labelled according to their material types and subtypes, which enables versatile analysis. For example, a spectrum reflected from a birch plank could be labelled into several material categories T, such as "organic matter", "plant", "wood", "cut wood", and "birch". Using multiple categories per one spectrum would enable more versatile material identification: for example, by using information gained from other wood data, the program can recognize that the given oak sample is very likely wood even if the source library has no oak data.

A need for altering band splitting structure may arise according to available hardware capacity and known material types, and therefore the software should be able to adjust the source data coordinates according to the changes. This can be done by storing the source data as both spectra and the N-dimensional coordinates, and the latter ones are re-calculated if the band splitting is altered. Using readily calculated coordinates ensure faster identification analysis. The program also allows

including new coordinate-based identification procedures only by inserting more probability calculation VI:s (virtual instruments) into the identification VI.

4.11 Identification case not treated by the software: spectral unmixing

A remarkable issue with the current identification software is that the software can recognize only relatively pure samples; the mixtures of different materials cannot be identified directly by our instrument. In other words, our instrument is directly uncapable of spectral unmixing, or identifying materials in mixed-material spectra.

There are studies and algorithms made for spectral unmixing, or identifying materials in a spectrum of mixed materials. Typically, these methods estimate the spectra of material mixtures as in-between spectra between pure material spectra, or spectral coordinates that are located between the coordinates of pure-material spectra. For example, two-material mixtures would have spectral coordinates located on the line between the spectral coordinates of the two pure materials. [11]

An indirect way to identify material mixtures is to identify materials that have highest probabilities. However, this does not work if the material mixture is made of materials that have very differently shaped spectra, and therefore the measurement results in a spectrum that resembles none of the pure material spectra. We focus on spectra that *mostly* contain certain material. As our instrument illuminates only a small area at once (diameter approximately 5 centimeters), it is reasonable to estimate that most interesting spectra are mostly one-material spectra.

4.12 Testing the identification algorithms

In order to evaluate and optimize the spectral identification algorithms, it is necessary to test them in various situations with different targets and libraries. The following factors affect the identification processes and their calculation times;

- Identification method
- Spectral band division
- Variety of data in the material type library
- Using the non-alpha library or the alpha library
- Identification acceptance (the threshold to suggest that there is no material of given type in the reference data library or that data contains too much noise)
- Properties of the identifiable measured spectrum:
  - Standard deviation of the library spectrum σ_j^L and its scaling parameter VP
  - Number of sample spectra used for averaging an identifiable spectrum
  - Normalization of the spectrum

All these values were tested according to various criteria. An automated spectral evaluation software was built in order to perform a large number of identification tests, which are introduced in the following subchapter. In order to evaluate different

identification methods and their variants, automatic tests were programmed into the software. This required finding numerical criteria for the performance of the methods. The same tests can be used for optimizing the identification methods and material type libraries.

An intuitive way to compare identification methods is to make them identify several test spectra, after which the software determines the ratio of the number of correct material guesses out of all non-rejected spectra (correct identification ratio). This way can also be used for evaluating the accuracy of unknown material identification – in other words, how well an identification algorithm can recognize if the identifiable spectrum does not match with any material represented in the material type library. The correct guess ratio is one of the most important factors in the method evaluation, ideally coupled with low rejection ratio.

Chang [6] has also introduced three other criteria to evaluate the quality of the identification methods: spectral discriminatory power, spectral discriminatory probability, and spectral discriminatory entropy. As their names imply, these three variables indicate the discriminatory qualities of the algorithms. Let L be the employed material type library (in our case, either a non-alpha library or an alpha library), m_{method} a spectral measure, $\mathbf{S_i^M}$ a spectrum to be identified, and $\mathbf{S_j^L}$ a spectrum in L.

Spectral discriminatory power implies the ability to differentiate between two spectra by using the given identification method, and it is defined as [6]

$$PW^{method}(\mathbf{S_1}, \mathbf{S_2}, \mathbf{S_R}) = max\left(\frac{m_{method}(\mathbf{S_1}, \mathbf{S_R})}{m_{method}(\mathbf{S_2}, \mathbf{S_R})}, \frac{m_{method}(\mathbf{S_2}, \mathbf{S_R})}{m_{method}(\mathbf{S_1}, \mathbf{S_R})}\right), \tag{12}$$

where $\mathbf{S_R}$ is a reference spectrum against which two spectra $\mathbf{S_1}$ and $\mathbf{S_2}$ are compared. The higher $PW^{method}(\mathbf{S_1}, \mathbf{S_2}, \mathbf{S_R})$ is, the better the identification method m_{method} is in telling apart the difference between $\mathbf{S_1}$ and $\mathbf{S_R}$ and the difference between $\mathbf{S_2}$ and $\mathbf{S_R}$. The value of $PW^{method}(\mathbf{S_1}, \mathbf{S_2}, \mathbf{S_R})$ does not depend on the material type library or the material categorization choice, but instead it depends on the identification method configurations, such as band choices and averaging methods. Because the SD power test may give values that are very close to 1, our studied value is actually

$$\begin{aligned} PW^{method}_{studied}(\mathbf{S_1}, \mathbf{S_2}, \mathbf{S_R}) &= PW^{method}(\mathbf{S_1}, \mathbf{S_2}, \mathbf{S_R}) - 1 \\ &= max\left(\frac{m_{method}(\mathbf{S_1}, \mathbf{S_R})}{m_{method}(\mathbf{S_2}, \mathbf{S_R})}, \frac{m_{method}(\mathbf{S_2}, \mathbf{S_R})}{m_{method}(\mathbf{S_1}, \mathbf{S_R})}\right) - 1. \end{aligned} \tag{13}$$

This helps studying the smaller test values, as the minimum value of PW^{method} is 1 anyway.

As is intuitively seen and later shown in results, the value of $PW^{method}(\mathbf{S_1}, \mathbf{S_2}, \mathbf{S_R})$ depends on the used reference spectrum $\mathbf{S_R}$. Chang used the spectrum of a plant that differed from other plant spectra he used. Instead, we use four different artificial reference spectra; an even spectrum with a constant value of 1 and three evenly divided four-part spectra with constant values of 1.5, 0.5, 1.5, and 0.5, or values of 1.25, 0.75, 1.25, and 0.75, or values of 1.75, 0.25, 1.75, and 0.25.

The spectral discriminatory probability for the algorithm $m(\cdot,\cdot)$ expresses the probability by which the identifiable spectrum resembles a certain material type T_k and not the others. It is defined as [6]

$$p^{method}_{S_i^M,L}(T_k) = \frac{m_{method}(\mathbf{S_i^M}, \mathbf{S_j^L})}{\sum_{j=1} m_{method}(\mathbf{S_i^M}, \mathbf{S_j^L})}, \quad j \in T_k. \tag{14}$$

In our case, m_{method} can also be used to imply the identification result over several spectra of one material type.

Our way to use the test slightly differs from how Chang used it, because we use several library spectra per material instead of a single spectrum. Instead of using spectral measures m_{method} between a test spectrum and a library spectrum, we use inversions of the identification certainty (1-P, 0 if $P > 1$). This is a simple way to express the proximity between the test spectrum and the material type cloud, and it also takes into account the variance parameters. The SD probability test uses the inverted guess certainty of the correct material guess as the dividend, no matter what is the actual primary guess. Because the inverted identification certainty of the incorrect primary guess is smaller than the one of the correct guess, the SD probability value gains automatically penalty as the correct guess is used as the dividend and the incorrect one as a part of the divider. On the other hand, the effect of the penalty is decreased if the identification certainty of the correct material is almost equal to the incorrect guess. If a test spectrum has several correct material categories, for example when using subcategories and supercategories simultaneously, the dividend is the inverted probability of the best correct guess.

Over all library spectra, the $p^{method}_{S_i^M,L}(T_k)$ values result in a probability vector

$$\mathbf{p^{method}_{S_i^M,L}} = (p^{method}_{S_i^M,L}(T_1), p^{method}_{S_i^M,L}(T_2), \cdots, p^{method}_{S_i^M,L}(T_K)). \tag{15}$$

This vector is called the spectral discriminatory probability vector of L with respect to $\mathbf{S_i^M}$. The smaller a single $p^{method}_{S_i^M,L}(T_k)$ value is, the better the measured spectrum $\mathbf{S_i^M}$ corresponds to material type T_k.

Spectral discriminatory entropy is an uncertainty measure of identifying $\mathbf{S_i^M}$ by using library L, largely based on the spectral information measure featured by Chang [6]. It is defined by using spectral discriminatory probabilities as

$$H^{method}(\mathbf{S_i^M}, L) = -\sum_{k=1}^{K} p^{method}_{\mathbf{S_i^M},L}(T_k) log\left(p^{method}_{\mathbf{S_i^M},L}(T_k)\right), \quad T_k \in L. \tag{16}$$

The smaller the entropy is, the more likely it is that the spectrum is identified.

5 Results

5.1 Identification test results

In order to evaluate the identification methods and their configurations, algorithm tests were run for the self-made spectral library (189 spectra, 41 material categories which include both subcategories and supercategories) and 25 test spectra to identify. Technically, the test spectra, stored into text files, are actually spectral time series, measured from a single target. The time series are used to form a single spectrum of the measured target by averaging them over time. From this onwards, "a test spectrum" can refer to both the time series and an averaged spectrum created out of it.

The time series of the test spectra were measured indoors, typically at a distance of 14 m, but six of them were measured further away at a distance of 40 m. The pump laser of the SC light source was used with maximum power (20 W, equaling approximately 16 W final SC power in the wavelength band at 1500–2200 nm), and illumination time was usually 100 ms when the samples were positioned 14 m away and 200 ms for the distance of 40 m; in any case, each measurement was optimized to maximize the quality of each single snapshot spectrum while avoiding saturating the spectrometer. In order to avoid the heat-up drift (see Chapter 3.2.2), the SC source was heated up and used for illumination for several minutes before recording the actual measurements.

Many of the test spectra were measured in the same situation as the respective library spectra, making them very similar and therefore easier to recognize. The test spectra files were marked with the correct material types and their supercategories, which were used for automatically checking the correctness of the material guess given by an identification method. While testing these spectra may not be an ideal indicator for the algorithm performance in real scanning situations, it offers a way to compare different algorithm configurations, as these configurations undergo the same tests anyway.

The test spectra and the categories originally included also metal samples, but they were excluded from the final test runs. This is because the reflection spectrum of non-roughened metal differs depending on the angle, and therefore it is more difficult for our setup to reliably identify those. We tried using roughened and therefore more diffusive metal samples, but as the created metal categories disturbed identifying other categories, the metal categories were omitted from the final library.

The algorithm tests were run by using three different material category divisions; either (1) using all material categories existing in the material dictionary including both subcategories (single materials, such as "lingonberry shrubs" or "cardboard") and supercategories (boarder material categories, such as "vegetation" and "paper product"), (2) using only subcategories, or (3) using only supercategories. In actual identification applications, subcategories are likely an option, as in them the material choices are limited differently. However, as mentioned in Chapter 4.10, using supercategories is beneficial when attempting to recognize materials that are not present in the library, such as identifying that oak is a tree even if the library has

no oak spectra.

The algorithm tests employed four different band configurations (detailed in Figure 23) and variance parameters in order to create as a large variety of different identification scenarios as possible, therefore testing the typical performance and limits of each method. A test spectrum was rejected if all identification results for every material type gave less than 5% certainty.

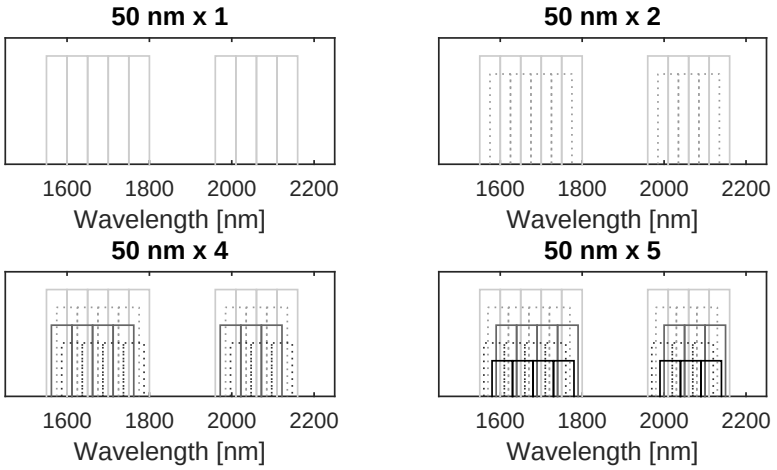


Figure 23: The band configurations used in the identification performance tests.

The variance parameters were mostly chosen from the semi-logarithmic sequence

$$\begin{aligned} \text{variance parameters} &= (1, 0.5, 0.25) \cdot 10^N, N \in \mathbb{Z} \\ &= (..., 1, 0.5, 0.25, 0.1, 0.05, 0.025, 0.01, 0.005, ...), \end{aligned} \tag{17}$$

or, if a denser parameter division was needed, the parameters were chosen from the sequence

$$\begin{aligned} \text{variance parameters} &= (1, 0.75, 0.5, 0.35, 0.25, 0.2, 0.15) \cdot 10^N, N \in \mathbb{Z} \\ &= (..., 1, 0.75, 0.5, 0.35, 0.25, 0.2, 0.15, 0.1, 0.075...). \end{aligned} \tag{18}$$

From these sequences, the variance parameters were chosen so that the test spectrum rejection ratio ranged from accepting all test spectra to rejecting them all. This kind of parameter choice enables seeing the full variability of the performance within an identification method. There is a tradeoff between rejection ratio and correct identification ratio; in the hoped case, the more test spectra are rejected, the larger part of the remaining samples are identified correctly.

Sometimes additional all-rejection variance parameters have been used in the tests. This is done for either of these two reasons: first, the test software requires having an equal number of variance parameters for both non-alpha and alpha library options of each configuration; second, additional points are needed for showing the scale of the variance parameter axes. These variance parameters were chosen preferably from the all-rejection end of the semi-logarithmic sequences, because the all-rejection configurations do not affect the other results.

5.1.1 Identification-rejection graphs

A crucial part of this analysis is comparing the correct identification ratio of each method configuration against their rejection ratio. The performance of a method can be evaluated by plotting these ratios on the same graph to see the nature of the tradeoff between correct identification ratio and rejection ratio. Since these graphs form a major part of the analysis in this thesis, they are featured early in this subchapter. From this onward, these graphs are called identification-rejection graphs.

Figure 28 shows an example of an identification-rejection graph. In these graphs, each dot represents a combination of calculated correct identification ratio (vertical axis) and rejection ratio (horizontal axis) that results from an identification method using a certain set of parameters, such as the averaging choices, the band division choice, and the variance parameters. The shade of the dot implies the number of the configurations that result in the given ratio combination. It is worth clarifying that the correct identification ratios are calculated *out of the samples that were not rejected*, and the figures exclude configurations that reject all test spectra. Since the number of identification ratio possibilities changes when the test spectra are rejected, the gray plus signs are used to represent unoccupied correct guess/rejection ratio locations.

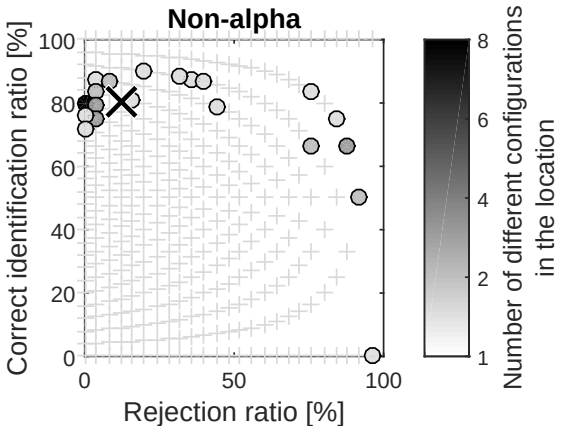


Figure 24: An example of an identification-rejection graph. The distance method was used with the non-alpha library and all material categories.

The ideal location of a configuration dot is the upper left corner of the graph, or the point where no test spectrum is rejected yet all spectra are correctly identified. No algorithm configuration manages to reach this optimal point, but the proximity to the ideal performance offers information about each algorithm configuration. Different methods can also be evaluated quantitatively by studying the shapes of the configuration dot clouds.

To compare the shapes of the configuration clouds, the following two numerical

values are calculated:

1. **average identification ratio at low levels**, which equals the average identification ratio at low rejection ratios (less than 25%), illustrated by a large black cross in the graphs; and

2. **the maximum correct identification ratio**, which equals the highest achieved identification ratio. We can also calculate the rejection ratio of the configuration with the highest correct identification ratio.

Now when the identification-rejection graphs are introduced, we apply them for evaluating the effects of averaging spectra over time series and effects of moving averages on the calibration spectra.

5.1.2 Effects of averaging over subsequent spectra and averaging the white reference spectrum

In order to limit the calculation effort and to clarify the graphs, an analysis was performed on two kinds of averaging effects: averaging a test spectrum from subsequent snapshot spectra in a test spectrum time series, and smoothing the white reference (WR) spectrum by using moving average. The averaging of the snapshot spectra was intended to remove noise from the average test spectrum, while the averaging of the WR spectrum is intended for making the WR spectrum more "neutral" and therefore to compensate the SC drifting.

The numbers of snapshot spectra used for the averaged test spectrum were 1, 3, 5, 7, 10, 15, 20, 50, and all available spectra. However, one must note that not all test spectra contained long time series, but the lengths of the time series varied between 8 and 549 spectra. The used channel averaging widths of the WR spectra were 1, 3, 5, 7, 11, and 15 channels, equaling wavelength bands of 6.6 nm, 19.8 nm, 33.0 nm, 46.2 nm, 72.6 nm, and 99 nm, respectively.

The averaging comparisons were done by limiting the spectral configurations into desired averaging values, drawing identification-rejection graphs out of them, and comparing graphs of different averaging types. After that, we calculated the average value of the correct identification ratios of the method configurations that resulted in a rejection ratio lower than 25%. Figure 25 shows examples of calculating the average identification ratios. The effect of the snapshot spectrum or white reference spectrum averaging can be seen by plotting the averaged correct identification rates against the used test snapshot spectrum or WR spectrum averaging values.

Figure 26 shows the effect of averaging different numbers of subsequent spectra on correct identification levels on low rejection levels. The results show varying effects of averaging numbers; almost all method configurations appear to peak at 5 snapshot spectra, after which they usually dive to lower values. When using the alpha library, increasing the number of averaged spectra generally increases the performance, albeit not notably compared to the number of required spectra.

These results align with the earlier Allan noise test results (Chapter 3.2.2); averaging clarifies the data to an extent, but then drifting effects worsen the performance.

However, it is notable that a typical averaging peak is located at 50 snapshot spectra (though since the number of averaged spectra was sparse in this area, the exact behavior of high averaging numbers is not clear). This is probably because averaging 50 snapshot spectra improves identification of some spectra and worsens that of some some, especially of those that benefitted from averaging 5 averaged spectra.

In realistic scanning situations, we prefer short illumination times, for which one reason is naturally the limitation of total scanning time. However, more importantly, the measured spectra become increasingly vulnerable to SC drifting when there is a larger time gap between measuring the white reference spectrum and the actual target spectrum. Because most methods show a peaked performance at 5 spectra, in later inspections we use that averaging number even if it is not optimal in every situation.

We studied similarly the effects of employing moving average on the WR spectra, but by varying channel widths instead of the number of averaged snapshot spectra. All tested spectra in this phase were averaged by using 5 snapshot spectra – other averaging values were also tried with similar results. Figure 27 shows the results of this analysis. The results indicate little difference between different WR averaging options. The graphs imply that this kind of averaging has little effect on the identification performance. In fact, non-averaged WR spectra appear to give better results than averaged ones.

Based on these results, in the band division and the variance parameter analysis we use time averaging of 5 spectra and non-averaged WR spectra, as the 5 spectra averaging gives peak results with reasonable number of snapshot spectra and averaging the calibration curve did not improve the results.

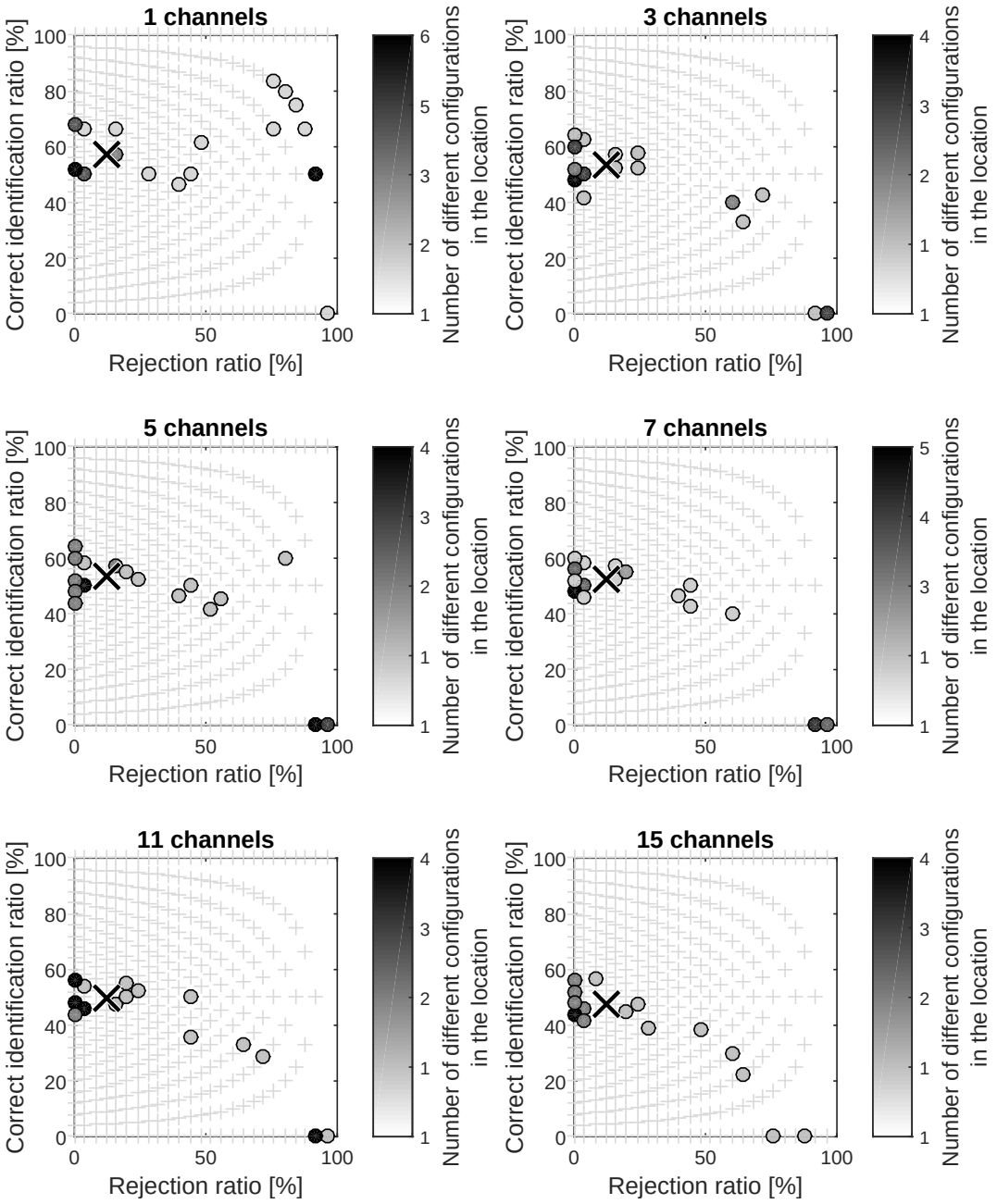


Figure 25: Effect of averaging WR spectra over spectral channels. In this case, the graphs use the distance method, alpha library, and subcategory division. The black crosses indicate the average value of the correct identification ratios of the method configurations that produce the rejection ratio equal or below 25%.

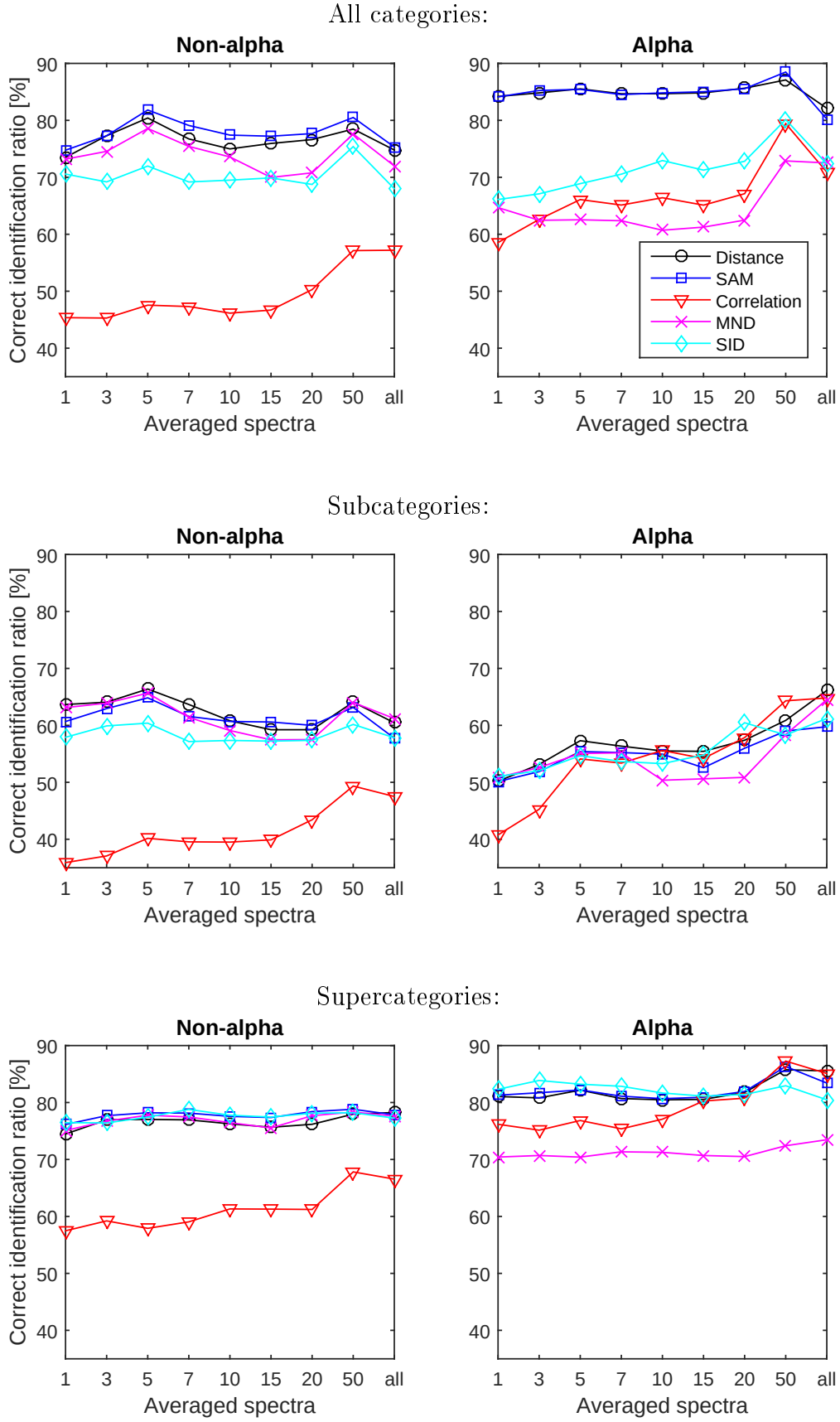


Figure 26: Effect of averaging subsequent snapshot spectra of test spectrum time series.

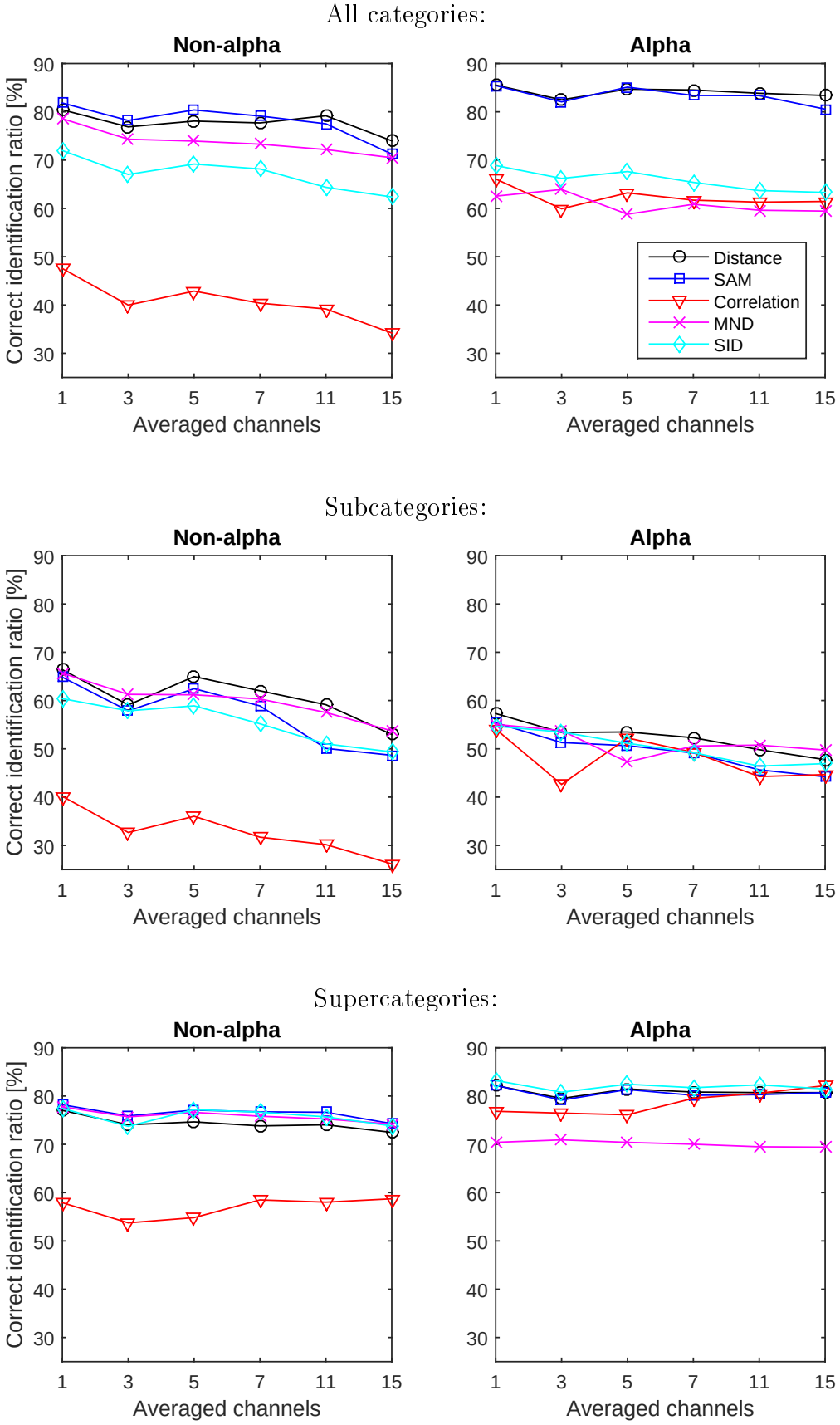


Figure 27: Effect of using moving average on the spectral channels of white reference spectra.

5.1.3 Shapes of identification-rejection graphs for each method

Appendix A shows identification-rejection graphs for each method with different band configurations and variance parameters, by averaging test spectra from 5 subsequent snapshot spectra and normalizing them with non-averaged WR spectra. Intuitively, it appears that decreasing the variance parameter causes the algorithm to reject more test spectra, or, in other words, to determine that the identification probability $P(T_k)$ for all materials T_k is less than 5%. Quite often but not always, the correct identification ratio increases as the variance parameter decreases, as a greater number of difficultly identifiable spectra are rejected. Therefore, the choice of an ideal variance parameter can be a tradeoff between the correct identification ratio and the rejection ratio: the overall identification results become more reliable when dubious spectra are rejected.

To study the differences of the configuration clouds, we calculated two ratio values to illustrate the optimal corners of the configuration clouds (see their definitions in Chapter 5.1.1). Tables 1–2 show these values for each algorithm configuration, alpha library choice, and material category choice combination. For example, as predicted, the supercategory division was best in identifying spectra, as the supercategories are the most board categories and therefore offer the easiest material guesses. Also all-category division offered good identification results, but subcategory division shows worse results. In the case of the subcategories, using alpha libraries could worsen the low-rejection identification results. Using alpha library instead of non-alpha library sometimes appears to improve the results and sometimes not, though statistical error may hide actual trends. Figure 28 illustrates typical behavior of most methods.

Generally, it appears that achieving perfectly reliable identification (or, correct identification ratio of 100%) is a difficult and sometimes even impossible task. For example, the distance method can never achieve the correct identification ratio of 100%, albeit it can come close (over 90%). This is shown in Table 1 and Figure 28. The other methods follow the same trend, except for the correlation method, which behaves overall very differently to the other methods. With low rejection ratios ($<$25%), the correlation method shows significantly lower correct identification ratios than the other methods, but as the rejection ratio increases, the correlation method provides perfect identification ratios faster than any other method. Figure 29 shows the identification-rejection graphs for the correlation method.

The identification-rejection graphs were also drawn separately for different band configurations, and their low-rejection identification values are shown in Figure 30. This helps seeing the effect of the band configurations on the identification performance. The graphs show that with low rejection ratios, the 1-fold band configurations usually give the best identification results. This is especially true when employing the alpha library and/or the subcategory division. However, the correlation method shows again different behavior here by preferring multi-fold band configurations over 1-fold ones. This is clearly caused by the original use of linear correlation; often, the more data is used for calculating linear correlation, the more certain estimates can be made of the correlation between the data sets.

Table 1: Average correct identification ratios and their standard deviations on low rejection ratios (≤25%) for different method configurations, calculated from graphs in Appendix A. Accuracy of 1% is used in these tables.

	All dictionary [%]		**Subcategories [%]**		**Supercategories [%]**	
	Non-alpha	Alpha	Non-alpha	Alpha	Non-alpha	Alpha
Distance	80 ± 4	86 ± 3	66 ± 6	57 ± 8	77 ± 4	82 ± 3
SAM	82 ± 4	85 ± 3	65 ± 5	55 ± 8	78 ± 4	82 ± 4
Correlation	48 ± 18	66 ± 10	40 ± 21	54 ± 7	58 ± 22	77 ± 4
MND	79 ± 4	63 ± 8	66 ± 7	55 ± 7	78 ± 3	70 ± 7
SID	72 ± 7	69 ± 6	60 ± 6	55 ± 6	77 ± 4	83 ± 3

Table 2: Highest identification ratios for different method configurations and their locations on the rejection ratio axis (in brackets), calculated from graphs in Appendix A. Accuracy of 1% is used in these tables.

	All dictionary [%]		**Subcategories [%]**		**Supercategories [%]**	
	Non-alpha	Alpha	Non-alpha	Alpha	Non-alpha	Alpha
Distance	90 (20)	93 (44)	75 (4)	83 (76)	90 (60)	88 (0)
SAM	92 (4)	100 (96)	72 (0)	100 (96)	88 (68)	100 (96)
Correlation	100 (76)	100 (68)	100 (80)	100 (68)	100 (68)	100 (68)
MND	88 (36)	77 (12)	76 (0)	75 (68)	86 (72)	88 (68)
SID	83 (28)	78 (64)	73 (56)	75 (68)	88 (68)	88 (0)

All identification methods except the correlation method show similarity in the shapes of the empty area in the ideal corner (no rejection, perfectly correct identification); it is easy to achieve quite high identification ratios by rejecting no spectra, but significant improvements in the identification ratio cannot happen without great sacrifices in the rejection ratio, and neither that is proven to work.

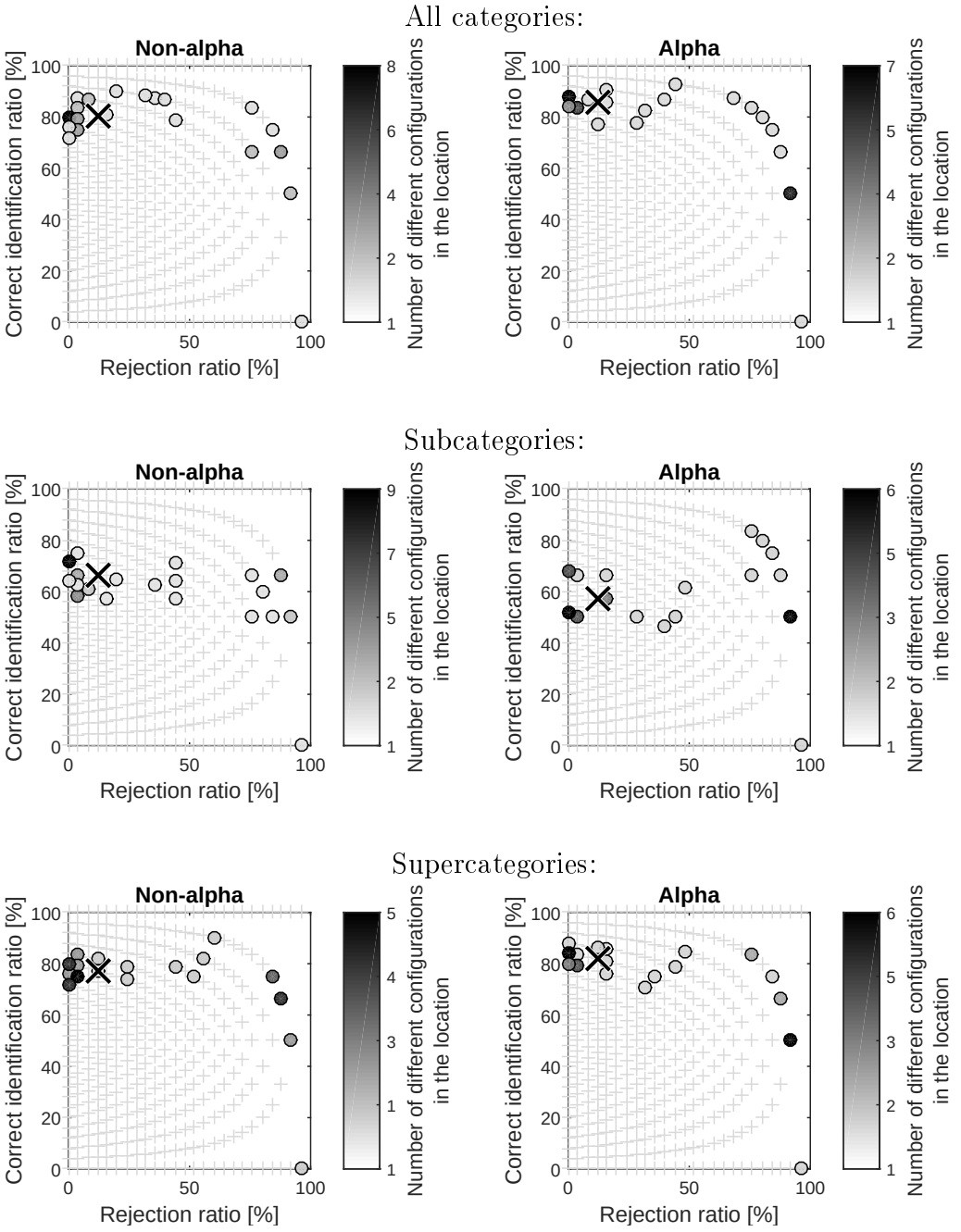


Figure 28: Rejection ratio versus correct identification ratio for the distance method, using all categories (upper row), only subcategories (middle row), or only supercategories (lower row), with the non-alpha library (left) or the alpha library (right).

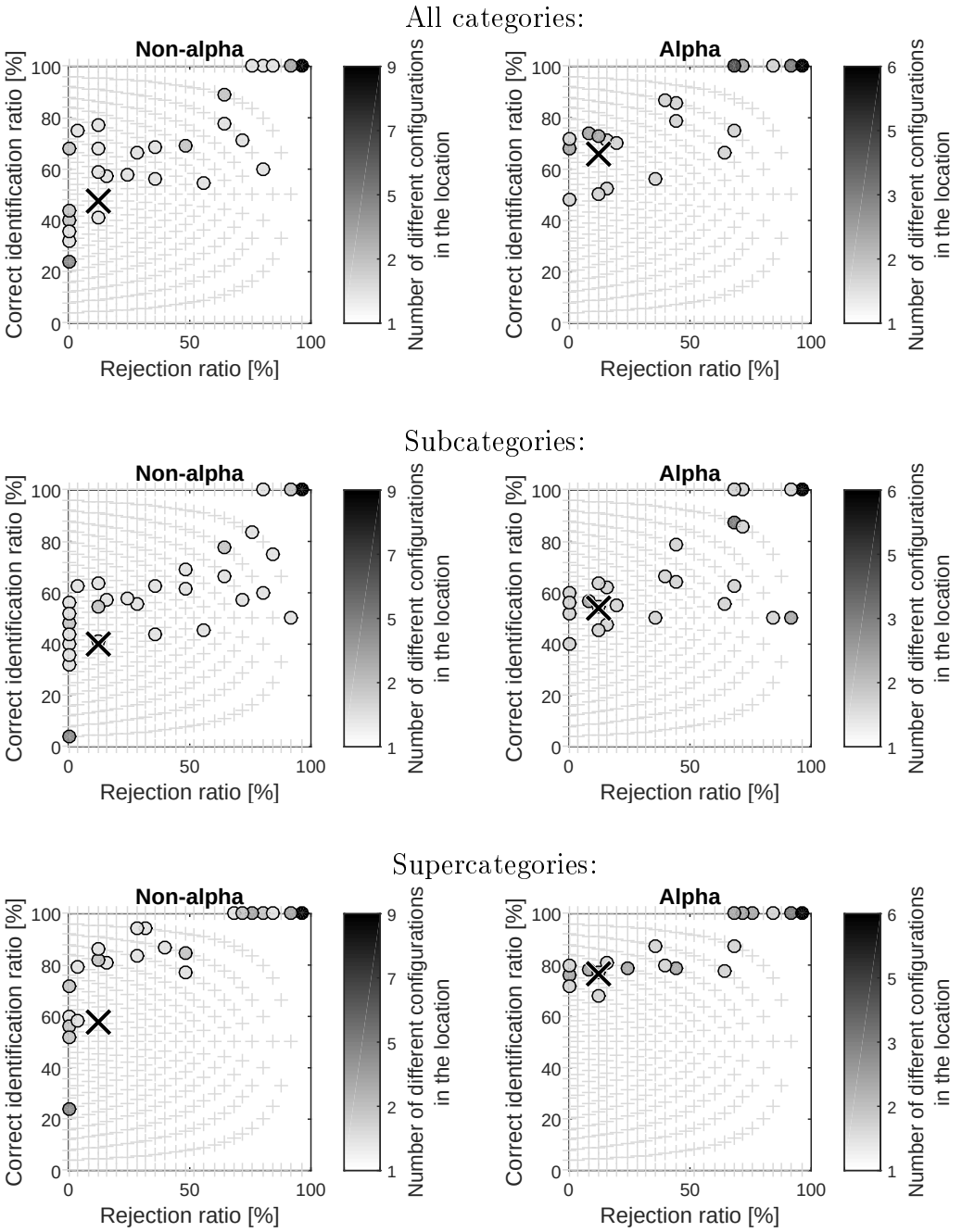


Figure 29: Rejection ratio versus correct identification ratio for the correlation method, using all categories (upper row), only subcategories (middle row), or only supercategories (lower row), with the non-alpha library (left) or the alpha library (right).

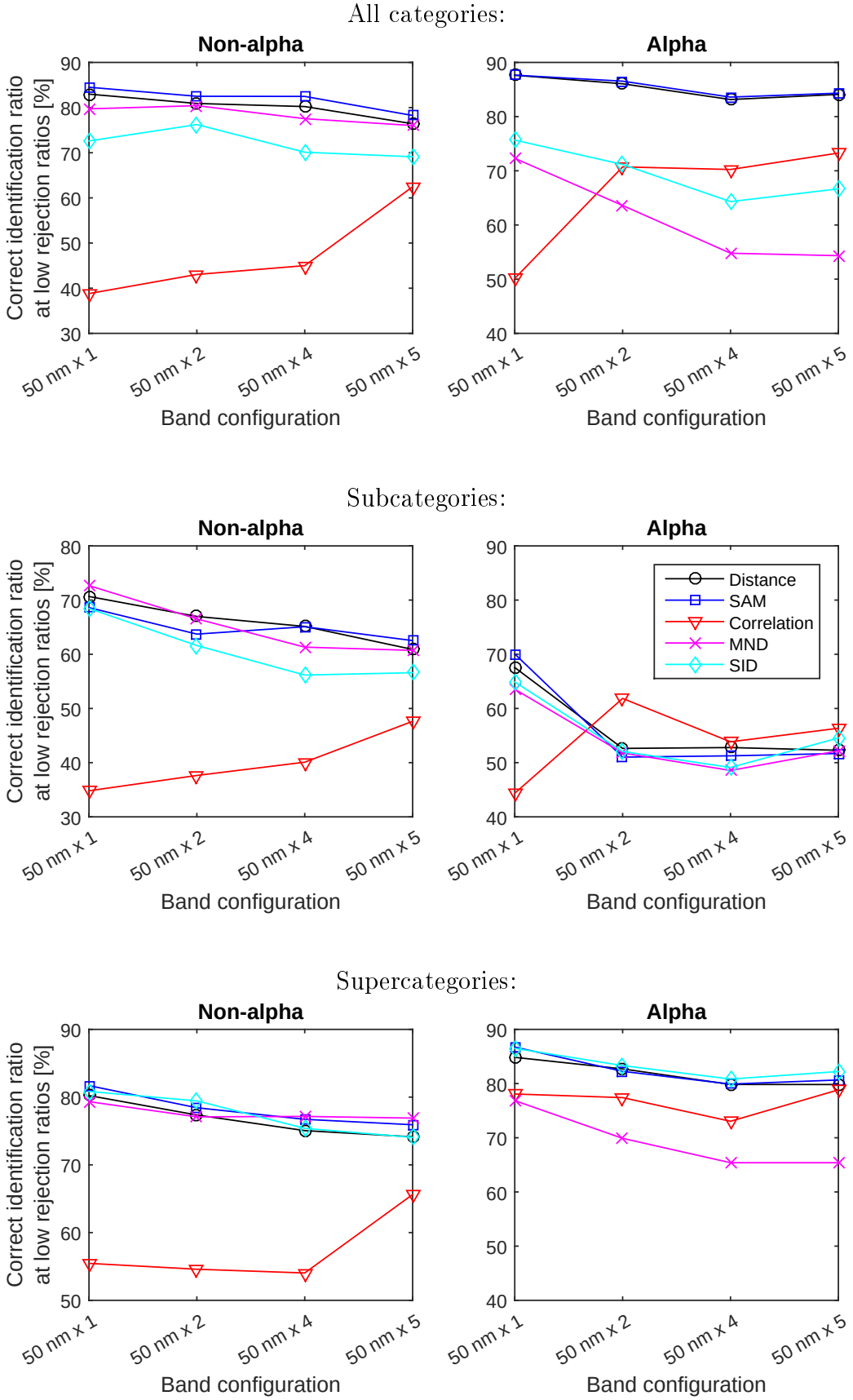


Figure 30: Effect of the band configuration choice on the average correct identification ratios of the configurations with the rejection ratio below 25%.

5.1.4 Band configuration and variance parameter graphs

Appendix B shows graphs that illustrate the effects of the band configurations and variance parameters on the correct identification and rejection ratios, and Figure 31 provides an example of these graphs. The dots indicate each band choice and variance parameter configuration, and the shade of the dot indicates the correct guess ratio or the rejection ratio (the darker, the greater). Crosses stand for configurations that reject all spectra, black diamonds express the correct identification ratios of 100%, and white diamonds express identification or rejection ratios of 0%. The diamond symbols help differentiating 100% and 0% values from the values close to 100% or 0%.

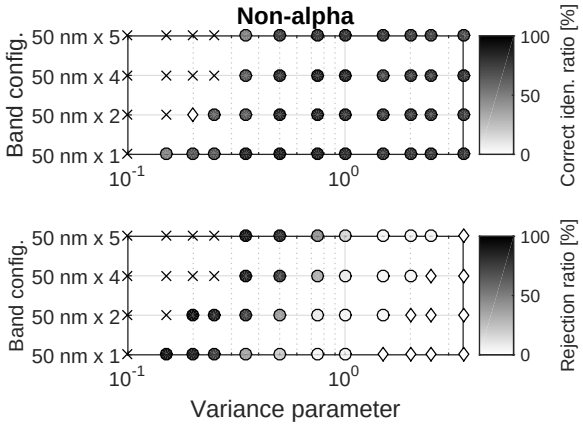


Figure 31: An example of band configuration choice and variance parameter choice table. In this table, the distance method was used with the non-alpha library and all material categories.

When seeking certain kind of identification and rejection ratios, the variance parameter does not appear to depend on the band configurations in the case of the SAM method and the SID method, but it has some effect on the distance, correlation, and MND methods. In the case of the latter ones, denser band configurations require larger variance parameters, and this shifts exponentially. Instead, material categorization (all categories versus subcategories versus supercategories) does not appear to have a remarkable effect on the choice of the band divisions and variance parameters, or at least it does not appear to affect the rejection ratio; it does have an effect on the identification ratio, as discussed in the previous chapter.

Different methods appear to favor different variance parameters; for example, the SAM method favors parameters ranging between 10^{-2}–10^{-1} (non-alpha) or 10^{-3}–10^{-2} (alpha), while the correlation method prefers magnitudes of 10^{-4}–10^{0} (non-alpha) or 10^{-6}–10^{-2} (alpha). This is caused by the different scales of the spectral measures of the methods; for example, the value of m_{SAM} can be virtually anything between 0–90 degrees, while $m_{correlation}$ ranges between 0–1.

The use of the non-alpha library or the alpha library clearly affects the scale of the preferred variance parameters; when using the alpha library instead of the non-alpha library, all methods need smaller variance parameters. However, this does

not actually tell about the superiority of either choice, since alpha library spectra have anyway greater variance than non-alpha spectra.

5.1.5 Differentiation capability and identification certainty

The method configurations were also tested according to their capability to differentiate spectra and to their identification certainty, using tests suggested by Chang [6] (Equations 12–16). For the sake of clarity, the analysis used only spectra with averaging over 5 spectra and no calibration spectrum averaging.

Spectral distribution power

First we used spectral distribution power (SD power) test, given by Equation 12 (or rather, Equation 13). SD power measures the differentiation capability of a spectral identification method by comparing two spectra against a reference spectrum. We used four different reference spectra, shown in Figure 32.

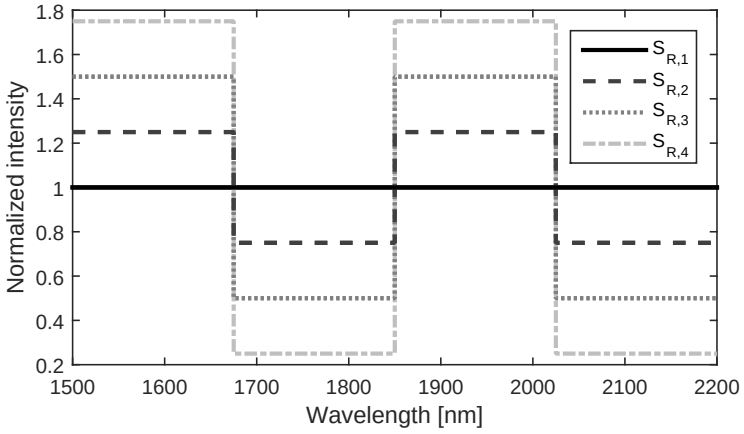


Figure 32: The reference spectra $S_{R,i}$ used in the SD power tests.

The material category distribution or the library choice does not affect the SD power measure, but instead the identification method settings have an effect. In fact, most methods are also independent from the variance parameters, because it does not affect the spectral measure m_{method}. The only exception is the MND method, as it gives directly the spectral density values, and therefore its differentiation capacity depends on the spread of its density clouds. Because the MND method gives no spectral measure directly, we use an artificial measure, which is expressed as

$$m_{MND,12} = 1 - \frac{p_{MND,12}}{max(p_{MND,12})}, \tag{19}$$

where $p_{MND,12}$ is calculated according to Equation 7. In other words, we determine the probability by which the spectra are *not* similar. This spectral measure functions similarly to the other spectra measures, indicating greater similarity when the measure is decreasing.

Because the comparisons involve all test spectra pair combinations, each identification method contains $(25^2 - 25)/2 = 300$ unique spectrum pairs, and therefore SD power gives actually 300 measures for each method configuration. Therefore, mean and median values values were chosen to study the identification method differences. The resulted mean and the median values are shown in Figure 33. The figure shows that the SID method shows often greater SD power values than other methods – except when the reference spectrum becomes less even, such as when using the spectra $S_{R,3}$ and $S_{R,4}$. This indicates that as such, the SD power test is not a simple indicator of the performance of the methods. The immense differences between the mean and median express the skewness of the SD power test results: especially SID appears to be prone to give extreme SD power values on occasion.

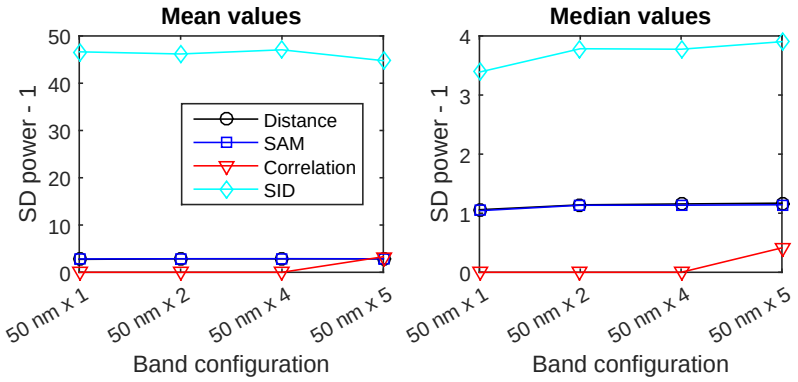
The distance and the SAM methods express similar differentiation capacity, which is not surprising: the distance method and the SAM method virtually calculate the same spectral value, with the relation of $m_{distance} \sim 2sin(m_{SAM}/2)$. In these SD power tests, the distance method shows slightly better performance than the SAM method, simply because the test spectra are not quite ideally scaled and therefore they look more different to the distance method than to the SAM method.

The MND method is not shown in these graphs, because it is dependent on the applied variance parameter. Therefore, the MND method must be studied separately by using tables featured in Figures 34. In these tests, we used the mean value of a reference spectrum as its standard deviation, because then its variance parameter directly expresses the magnitude of the ideal relative standard deviation. We notice that the SD power values for the MND method increase when the variance parameters increase; this may be a direct result from the definition of the artificial spectral measure for the method, m_{MND}. When using variance parameters around 10^0 and the reference spectrum $S_{R,1}$, the MND method gives comparable SD power values to the SAM and the distance methods, but the SD values rapidly decrease when less even reference spectra are used.

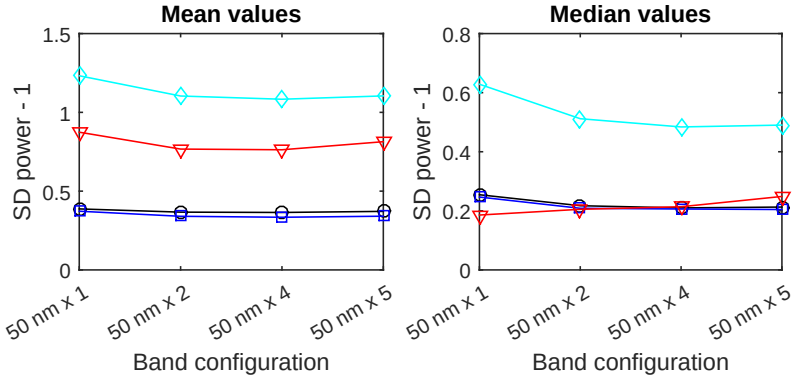
The band division appears to have an effect on the SD power, even though the behavior is not quite uniform. In most cases, however, adding more bands of the same width does not appear to increase the spectral discrimination but rather to *decrease* it. This appears initially surprising, given that adding more bands should mean giving each method more parameters to differentiate. Indeed, this is the case, but it is likely that the differentiation value decreases because the methods are given more bands that are similar in different spectra. Therefore, this shows the importance of picking up the spectra that actually show relevant differences.

A noteworthy issue with the spectral comparisons is that different methods have different ranges for their spectral measures. For example, the angles calculated by the SAM method can vary only between 0–90 degrees, and the correlation method varies only between the values 0 and 1. The distance and the MND method technically could have infinitely large spectral measures, but in reality they are limited because the compared spectra are normalized. The spectral measure of the SID method can technically be limitless, but it is limited by the behavior of the logarithmic function and the spectral normalization.

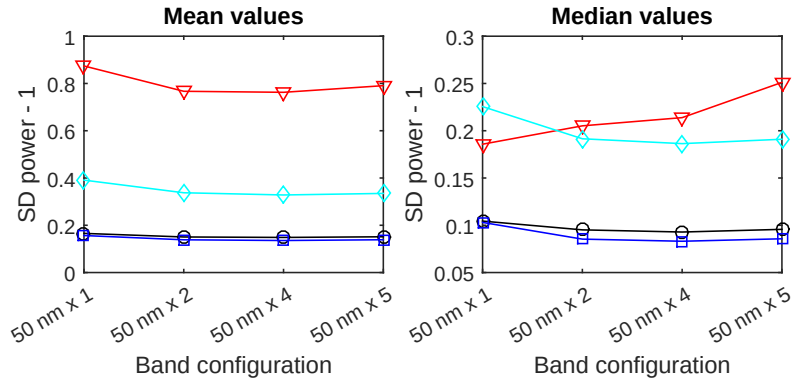
Reference spectrum $S_{R,1}$



Reference spectrum $S_{R,2}$



Reference spectrum $S_{R,3}$



Reference spectrum $S_{R,4}$

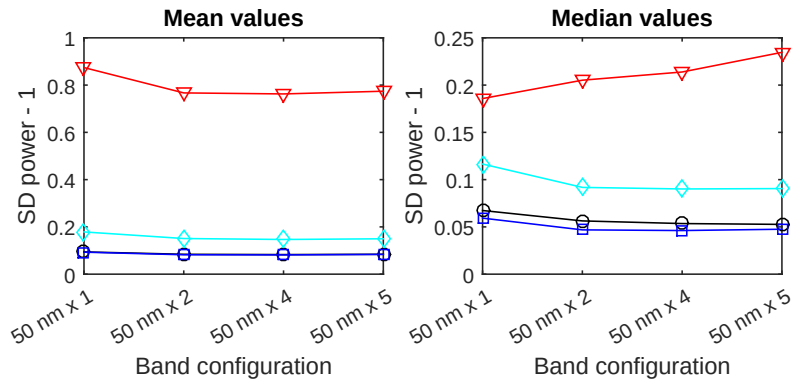


Figure 33: SD power values when using different reference spectra $S_{R,i}$.

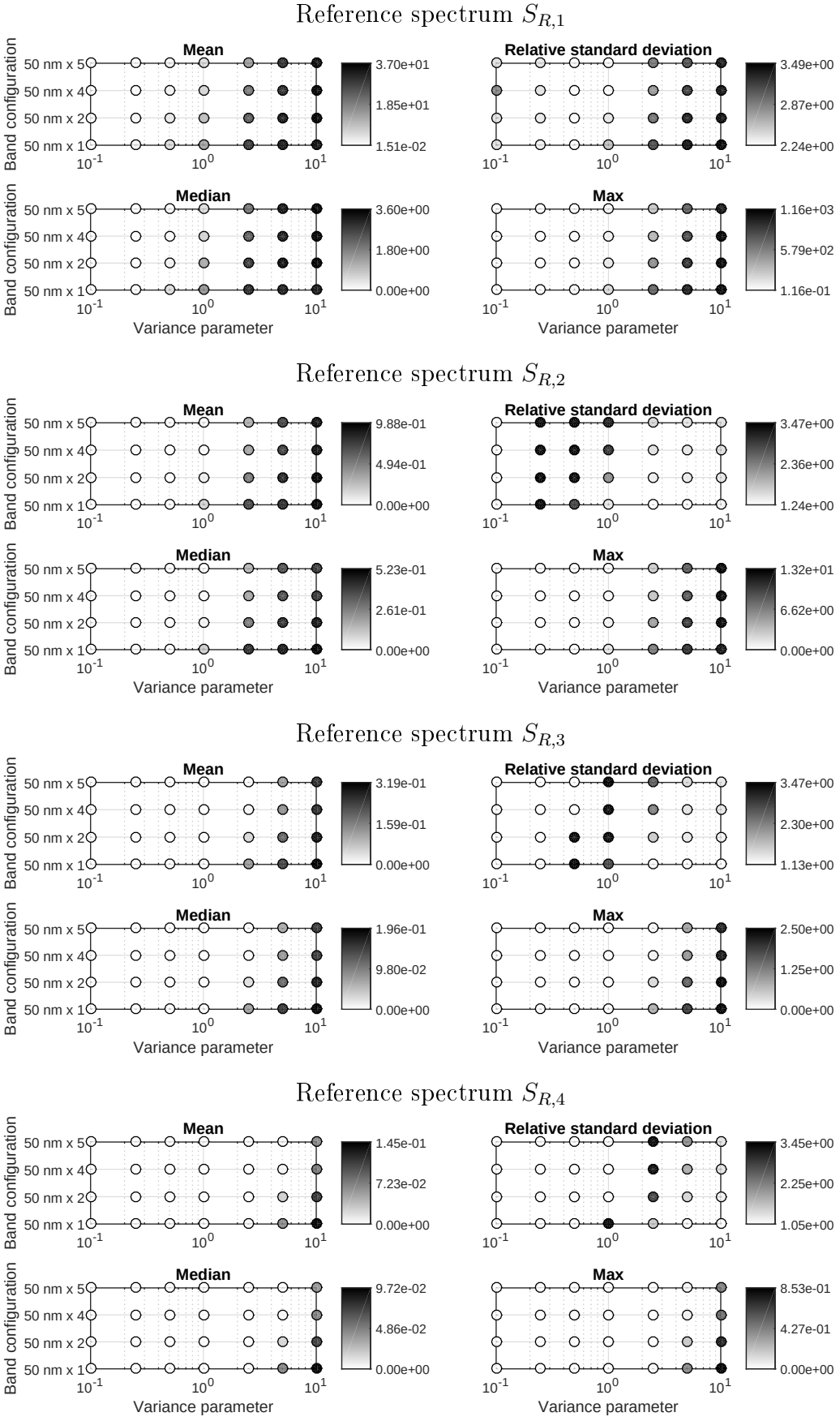


Figure 34: The SD power values for the MND method with different reference spectra $S_{R,i}$.

According to these tests, the SID method mostly shows the greatest differentiation capacity, only rivalled by the correlation method in the cases in which the reference spectrum is very uneven. Adding several bands usually weakens the differentiation capacity. For target finding use, a potentially useful reference spectrum could be a spectrum that is abundant in the scanning area but not interesting to the users. These spectra could be for example the spectra of most common plants in the field. Due to the varying behaviors and magnitudes of the spectral measures m_{method}, in addition to the fact that the MND method has no natural spectral measure, caution is needed when analyzing the SD power test results when comparing methods.

Spectral distribution probability

The test employed next was the spectral distribution probability (SD probability) test, given by Equation 14. This test expresses how certain the given identification method is about its guess, based on the material type library given to it. The smaller the SD probability value is, the better is the method configuration. Because each method identifies 25 test spectra, the methods also result in 25 SD probability values. Therefore, the SD probability results are shown as group values similarly to the SD power results; the only difference is that the minimum value is examined instead of the maximum value, as a smaller SD probability value indicates better performance. The tests were performed only for the subcategory division, as in the ideal case, the subcategory division is the one used for actual spectral identification.

The resulting SD probability tables are quite uniform; the larger the variance parameters are, the smaller are the SD probability values. This is slightly surprising, as it would be logical to assume extreme differentiation when the variance parameters are small. This is most likely because probability certainties of correct material guesses are small when using small variance parameters; this causes our "spectral measures" based on the material identification certainties to be large, too, and the divider of the SD probability equation does not compensate it similarly than with the large variance parameters. With the large variance parameters, the "spectral measures" appear to be proportionally larger than with the small parameters, resulting in smaller SD probability values.

All the methods have a similar magnitude of median SD probabilities, ranging approximately between $3.0{\cdot}10^{-2}$–$3.4{\cdot}10^{-2}$. Their mean values are also of similar magnitudes, but the SID method has lower mean values, ranging between $2.4{\cdot}10^{-2}$–$2.7{\cdot}10^{-2}$. Therefore, the SID method is once again the best algorithm what comes to its differentiation capacity.

Because all methods show very similar SD probability behavior and there are no remarkable differences between the non-alpha and the alpha library configurations, only a single SD probability result table is shown, in Figure 35. Despite of the fact that these tests did not offer information for these studies, SD probability could be useful when developing spectral libraries and for example material type categorization choices. We did not perform a categorization comparison in this

study, because the categorizations have very different numbers of material types, and therefore, the magnitudes of the SD probability values would be inherently different in each analysis.

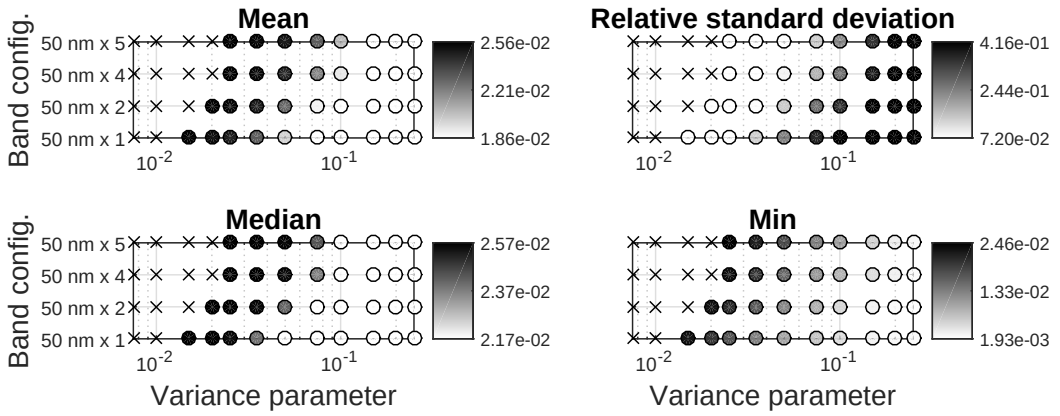


Figure 35: An example of SD probability analysis, performed by using the distance method and the alpha library.

Spectral distribution entropy

The next test employed was the spectral distribution entropy (SD entropy) test, given by Equation 16 as a direct derivative from the SD probability test. Because SD entropy is calculated for each test spectrum, each identification method configuration gives 25 values, and group analysis is once again needed. The results show similar traits than with the SD probability analysis, and we show no more tables except one in Figure 36.

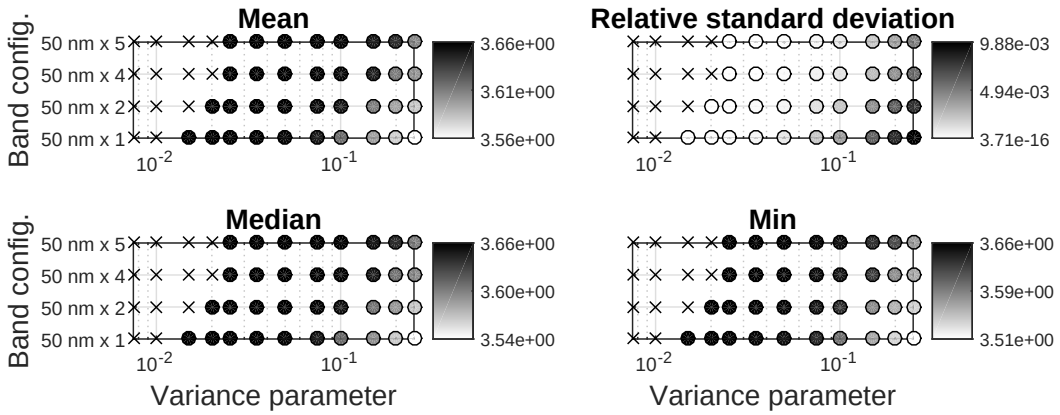


Figure 36: An example of SD entropy analysis, performed by using the distance method and the alpha library.

5.1.6 Calculation time

The calculation time of each method was estimated by using the clock of the computer, by measuring the time needed to run the identifications of all 25 test spectra and then dividing the time by the number of the test spectra. These measurements were performed with a freshly started Windows 7 PC while avoiding using other software. This kind of examination does not give accurate time estimates, but it helps seeing differences between the calculation times.

With our reference libraries and material categorizations, the identification calculation times for most methods were less than 1 ms when the computer is not otherwise heavily burdened. The exception is the correlation method, which required 10–40 ms for an identification. Admittedly, this is still a very short time compared to the scanner movement time, but it will become more of an issue when the number of reference spectra increases. However, data file handling and spectrum measurement times bring limitations to the speed of identification even if the calculation itself was fast. The file handling affects mainly when uploading the reference data and re-calculating the band divisions, which is not a major concern in this project.

5.1.7 Noise effects

This section presents a brief examination about the effect of the noise on the identification correctness. Here, the noise of a test spectrum S_i is defined as the average of the standard deviations of the two brightest areas of the spectra, 1650–1750 nm and 2000–2100 nm. Before calculating the standard deviations, the general trend of the spectrum was removed by using a moving average with a window size of 50 nm. In other words, the noise is defined as

$$\begin{aligned} noise &= mean\left(std(S_{i,no\ trend,1650\text{–}1750nm}), std(S_{i,no\ trend,2000\text{–}2100nm})\right) \\ S_{i,no\ trend} &= S_i - moving\ average(S_i, 50nm). \end{aligned} \tag{20}$$

The spectrum is automatically rejected if this noise value is greater than 1, or the standard deviation exceeds the typical mean value of the normalized spectra. This was not the case with any of the used test spectra, however.

We hoped to use this noise value to determine a limit for the acceptance of the identification so that the identification would require greater certainty when the noise is high. This was done by choosing a method configuration and plotting their noise and identification certainties onto a graph, marking correct guesses with crosses and incorrect ones with circles. An example of such a graph is shown in Figure 37. Several other graphs were made too, but we show only one picture of them as this examination was not successful.

Unlike hoped, the graphs did not show definitive limits between correct and incorrect guesses, even though in many pictures it seemed that increasing the noise decreases the overall identification certainty. However, the division between the correct and incorrect identifications by noise values was relatively even, making it

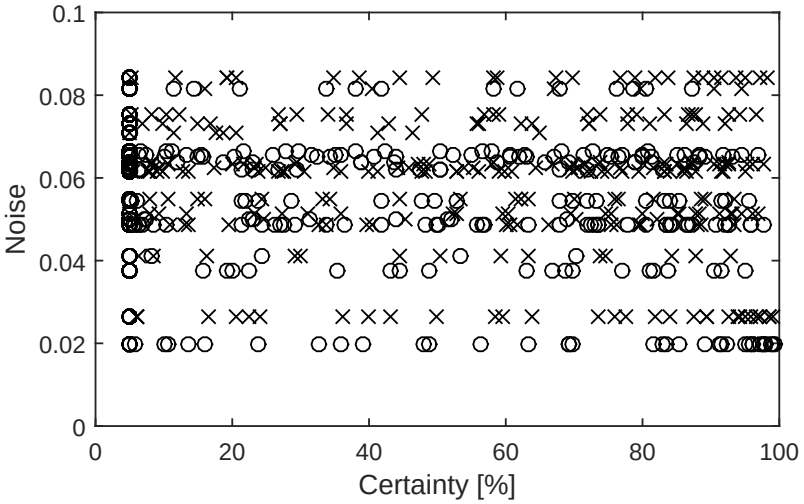


Figure 37: Noise values of test spectra plotted against their primary material guess certainties, using different MND method configurations with the alpha library and the subcategory division. The gap between the certainties of 0% and 5% is empty, because our software rejects identifications with a certainty of less than 5%.

impossible to use the noise values to determine the useful identification level. Due to the lack of the definite noise acceptance limit, we settled to reject identifications that give less than 5% certainty to the material guess. Alternative noise limit values were not tested due to time restrictions.

5.2 Scanner tests

In addition to single-spectrum measurements, also qualitative analysis was performed on the scanning identification performance. This was made by scanning an area and identifying the pixels afterwards by the aforementioned methods so that the variance parameters were adjusted until the best performance was reached. We used two-fold band division all times, as the band division showed mostly difference with differentiation capacities (fewer bands better), but it did not appear to have a great effect on the identification itself. The identification procedure is not yet implemented to work simultaneously with the scanning protocol, as the optimal identification protocol is not precisely established for now.

Due to SC drifting, the certainty of the identification process becomes doubtful as the calibration curve was measured typically only once right before the scanning. Scanning the spectra outdoors also sets numerous challenges to the spectral identification; varying weather conditions make spectra more difficult to identify, as the air temperature and moisture affect the propagation of radiation.

5.2.1 Indoors measurements

Figure 38 shows one of the test setups, scanned indoors at a distance of 14 meters. The setup contained the white reference, brown and red cardboard boxes, a sheet of grey foam rubber, and a steel sheet with three quarters painted with car paint and one unpainted. The scanned area also included a small peek to the corridor in order to test if the software is capable of ignoring too unclear (noisy) spectra.

After scanning, the software was qualitatively adjusted so that it would identify the targets as well as possible by each of the five identification methods and two library choices. The optimization focused on identifying large material areas as correctly as possible. Some analysis is also dedicated to the noise levels and intensity of the measured pixel spectra. We also studied how well the program can identify single material types, by choosing three most interesting material types and drawing red-green-blue (RGB) pictures out of the scan. This was done by converting the probabilities of the three interesting materials into the values of red, green, or blue hue of each pixel.

The analysis showed that with our library, different methods excelled in identifying different types of materials: for example, the SAM method using alpha library was well capable of differentiating foam rubber from the other targets, but it interpreted the white reference as a sample piece of granite. Other methods made similar mistakes between different materials. The non-alpha-library SID method was best in telling apart materials in this setting; it was the only method that was able to tell apart the painted plate, the foam rubber, the cardboard boxes, and the white reference. In this setting, the ideal variance parameter for it was 0.025. The optimized alpha-library counterpart had almost equal performance, except it easily mistook the white reference into a sample rock, though it was also better in identifying foam rubber. Figure 39 shows the identification results using the SID method with supercategory division and the non-alpha library.

We also ran the same tests while including metal categories. This confused the identification process, as the plastic categories were mistaken with those of metal, therefore giving plenty of incorrect guesses. As earlier discussed, the library has a great effect on the identification process, and poorly defined categories may ruin the identification process. Since metal is difficult to identify due to its high angle-dependent reflection, removing the metal categories from the library improved the identification performance.

In addition to the colored primary material guess graphs, there are also other ways to illustrate the spectral properties of the target setup, which could be used to assist identification. Figure 40 contains pictures of the signal intensity, noise (defined in Equation 20) and RGB identification graphs. In the intensity and noise graphs, white stands for high values and black for low values; in the RGB picture, red stands for white reference, green stands for paper products or dry vegetation, and blue stands for Type 1 plastic (which is also the supercategory of the used foam rubber).

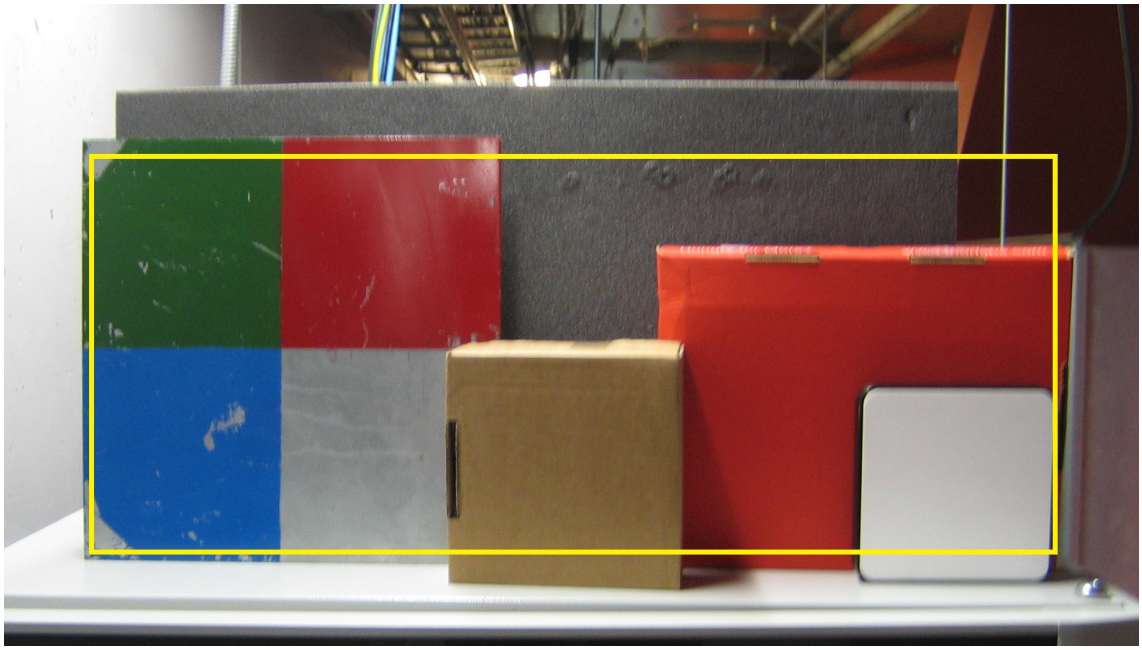


Figure 38: The indoors target setup. The yellow box indicates the scanning area.

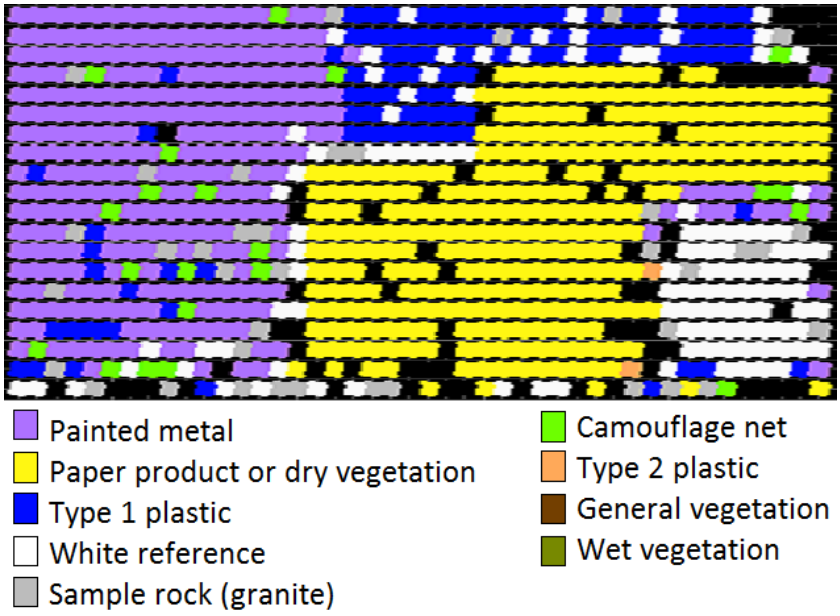


Figure 39: Identification image of the indoor target setup. In this picture, identification was performed by using the SID method with non-alpha library, with variance parameter of 0.0 and with a two-fold 50 nm band division.

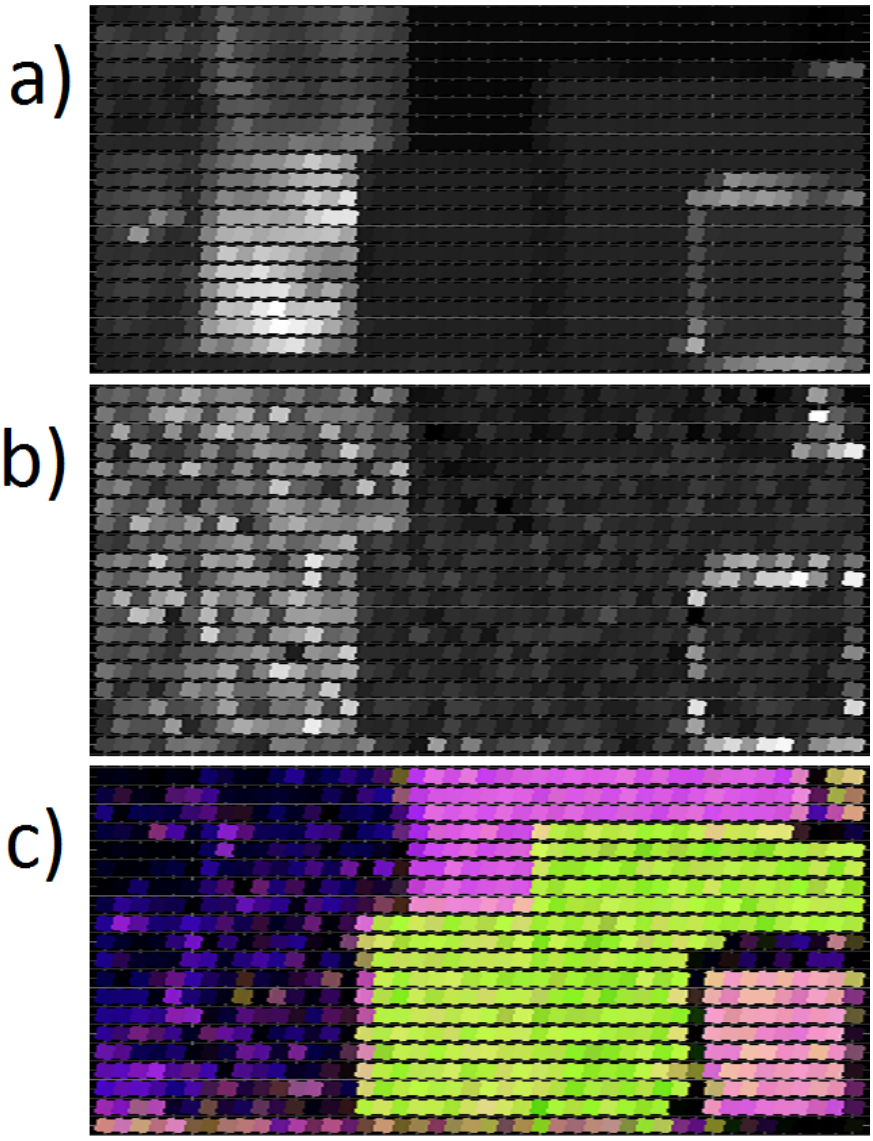


Figure 40: Intensity (a), noise (b), and RGB identification (c) data of the indoors target scan. In the intensity and noise graphs, white refers to high intensity/noise and black to low intensity/noise.

The intensity data shows that foam rubber and cardboard boxes reflect light more weakly than unpainted steel and the reflective black borders of the white reference. In addition to the "guess the best material for each pixel" protocol, another useful way to identify targets in the field is to use intensity data, as any artificial targets (such as metal objects) are highly reflective and most natural targets (such as plants) are not. The low reflection intensity also indicates weak backscattered signal intensity, and therefore it may decrease the signal-to-noise ratio, depending on the corresponding noise value.

Meanwhile, the noise graph could be used for evaluating the reliability of the identification results represented by the colored primary guess map. For example,

there are low intensity and high noise at the areas of the peek to the corridor at the upper right corner of the scanning area, which means that the scanned spectra here have poor signal-to-noise ratio and therefore it might be best not to try to identify them.

In addition to visualizing the primary material guesses by color dots, the software can identify the probabilities of 1–3 most interesting material types and represent them as a probability map, using RGB colors. This could decrease potential library errors related to single materials, as the material identification probabilities $P(T_k)$ do not depend on the rivalling material types. Instead, the order of the identification guesses is dependent on the overall library. A limitation in the RGB approach is the limited number of representable materials on a single picture, in addition to the effect of color sight of the graph interpreters.

5.2.2 Outdoors measurements

We also performed a test scan outdoors, by scattering items onto a small hill approximately 15 meters away. The outdoors target setup is shown in Figure 41. The featured area was scanned in two halves in order to decrease the effect of the SC drifting. By using the optimal non-alpha SID settings, the software was able to identify targets in the way that Figure 42 shows. This picture slightly differs from the setup photograph, since the photograph was taken from the right side of the lidar device. Most of the scan involved wet vegetation, which is visible in the primary guess color picture, but the software was less good in detecting targets set into the field. Being able to identify pixels that indicate the background is important ("We can be sure that at least here there is no tank"), but it is even more beneficial to be able to find materials of interest (such as the aforementioned tank).

Figure 41: The outdoors target setup. The yellow box indicates the scanning area.

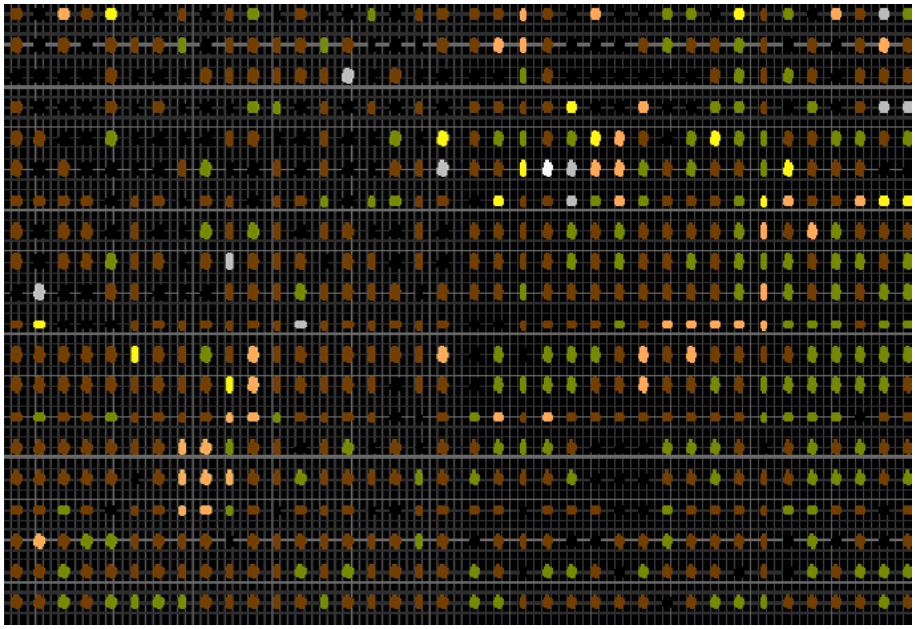


Figure 42: Identification image of the outdoors target setup, using the same settings than with the indoors setup. This picture is merged from two scan pictures with separate calibration curves. The colors of the identified materials are the same as those ones used in the indoors setup scan.

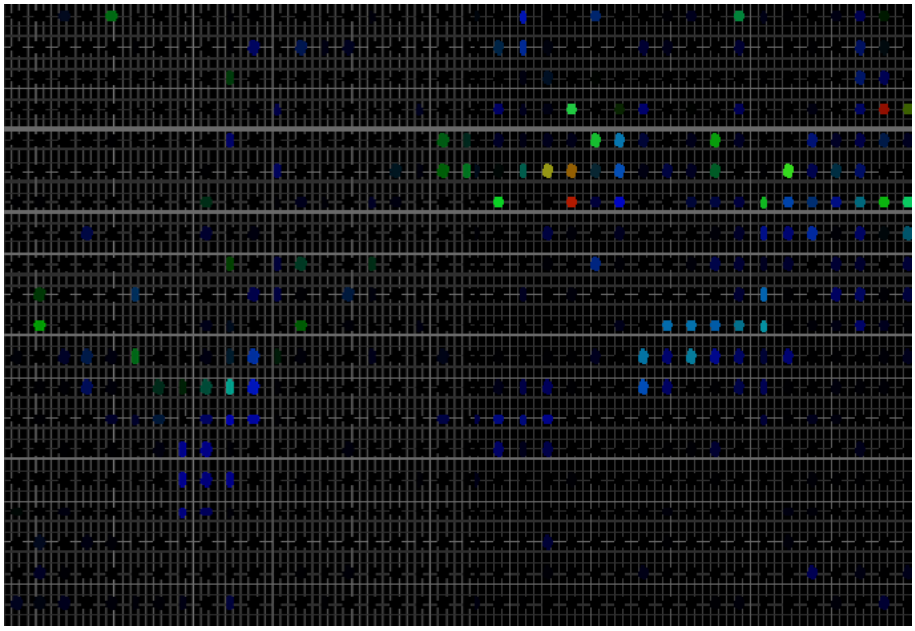


Figure 43: RGB identification image of the outdoors target setup. Red and blue stand for different plastic types, while green stands for paper products or dry vegetation.

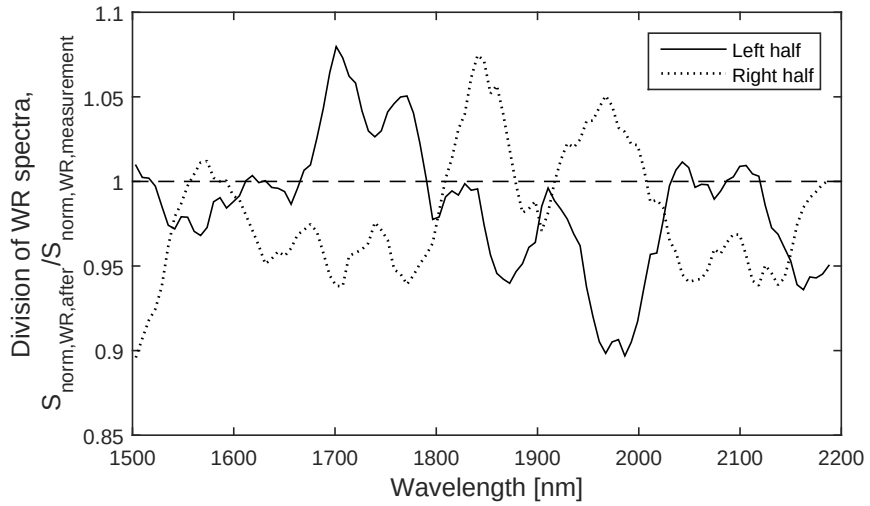


Figure 44: Examination of the WR spectra used in the outdoors scan.

In this case, the requested RGB approach functions better than the primary guess approach, as shown in Figure 43. Here red and blue are assigned for two plastic supercategories, while green stands for the paper products/dry vegetation supercategory. With the RGB approach, the plastic objects can be detected somewhat better, even though some targets are still missing and the software attempts to suggest the potential presence of plastic in the pictured Scots pine.

Since measuring each half scan took 14 minutes, the validity of the WR spectrum certainly changes over time, especially when taking into account weather changes. The WR spectra validity was studied by comparing the WR spectra before and after the scan, as shown in Figure 44. The comparison is performed by dividing the WR spectrum measured after scanning by the calibration spectrum used for the scanned spectra and then calculating a moving average with a window length of 50 nm over the division curve. As we suppose that the SC drifting makes spectra increasingly different over time, these comparisons of the WR spectra give an estimate for the maximum relative error in a 14 min measurement. Using a moving average helps us to visualize the effect of the employed 50 nm bands used in the spectral identification.

In Figure 44, we see that the division curves indicate even 10% relative errors; compared to typical material spectra (see Figure 18), this is a notable error, and it is safe to assume that the SC drifting results in a worsening identification in the last spectra of a single scan. This may also explain why the software tries to suggest that there is plastic in the pine tree; the scans are performed from down to up, and therefore the upper-part spectra have worse quality. This large error also explains why increasing the number of bands typically worsen the spectral identification results; using an increasing number of bands causes an increasing cumulative error.

As stated earlier, for now the calculation time of the identification process is insignificant compared to the time required for the scanning steps. Overall, the tests show that while finding the correct settings is not straightforward, the instrument can perform material identifications at least in ideal conditions.

6 Conclusions

A spectral identification software was built by using LabVIEW programming for the lidar instrument built at the Centre for Metrology and Accreditation (Mittatekniikan keskus, MIKES) in 2014. The further analysis of the results extended to the first months of 2015, and it was performed in MIKES Metrology at VTT Technical Research Centre of Finland Ltd., which MIKES merged with at the beginning of 2015. Five spectral matching methods were tested for the identification software: these methods were the distance, the spectral angle mapper (SAM), the correlation, the multidimensional normal distribution (MND), and the spectral information divergence (SID) methods. The methods were used to define material probability clouds by using multiple library spectra for each material, and these clouds were used for identifying the target material of the measured spectra. The library spectra of the same material could be used as separate spectra or be averaged together, former option being called a non-alpha library and the latter alpha library. A variety of parameters and tests were employed to investigate the performance of these methods in order to identify those that would be the most promising in the final version of the software. The current version of the software was used to run the performance tests.

The results express the existence of functional identification configurations for the lidar instrument, yet there are obstacles to overcome and alternative solutions to explore. With the current setup in an easy test environment, a moderately correct identification level (60–85% depending on the settings) can be achieved without rejecting spectra, while achieving higher identification levels requires high level of spectrum rejections, which is often impossible. However, high identification levels can be achieved with moderate rejection ratios, and some method configurations provided the best identification ratios at zero rejection levels.

Out of the attempted methods, all showed generally equal performance, having similar kind of identification-rejection configurations. The only exception was the correlation method, which showed poorer identification performance at low rejection rates but achieved perfect identification ratio at lower rejection ratios than the other methods. However, the correlation method also required multiple times longer calculation time compared to the others; with our library, the correlation method required even up to 10–40 ms for the identification compared to the typical values of less than 1 ms. One must notice that the identification performance depends on the employed library, and therefore these results cannot be generalized.

In military applications, reliable identification is preferred over minimizing the number of unidentified spectra; for example, the detection of an enemy tank could be used to launch a counter operation, and therefore the risk of false alarms should be minimized. This increases the importance of maximizing the correct identification, even if it meant sacrifices in the rejection ratio.

For most methods, the 1-fold band configurations often gave better results than the multi-fold configurations, with the exception of the correlation method that benefitted from the multi-fold configurations. The 1-fold band preference was especially visible when using subcategory division and/or the alpha library and when

the rejection rates were low.

It was found that the intensity-dependent distance method and the MND method were not entirely comparable to the other measurements due to the used normalization; our current setup normalizes the spectra by dividing their integrated intensity values between wavelengths of 1500–2200 nm. However, this band includes also water absorption bands, which makes the spectra less comparable when humidity conditions change. As the scaling has linear effect on the shape of the spectra, it has no effect on the SAM, the correlation, and the SID methods. If future work is directed to study further the distance and the MND methods, it is advisable to find an alternative way to normalize the spectra. This can be done by narrowing the integration wavelength bands to ones that are less affected by water absorption, but it increases the risk to lose some of the useful information.

The importance of the optimal band division was further shown in our spectral distribution power (SD power) tests; increasing the number of the bands does not necessarily improve the differentiation capacity of the method, but instead, it may even worsen it. This occurs because several bands most likely also contain bands with little variation, therefore making the spectra appear more "similar" to each other. This result also supports the band configuration preferences shown with the identification-rejection graphs. For specified future applications, it is vital to identify the bands of greatest importance which can be done for example by measuring typical spectra for a future use and then measuring the channel-wise deviations between the spectra.

When using the SD power measure in our differentiation tests, we noticed that the performance of the tests depends on the used reference spectrum. For example, the SID method was clearly the superior method when using an even reference spectrum, but the correlation method showed best performance for very uneven reference spectra. However, excluding the correlation method, the other methods remain in the same order despite the choice of the reference spectrum. Nevertheless, the effect of the reference spectrum choice should be taken into account when using this test in the future. The distance method and the SAM method showed similar behavior that was inferior to the SID method. MND method resulted in quite equal SD power values compared to the distance and the SAM method, even though its performance was highly dependent on the used variance parameter. However, one must note that the MND method used an artificial spectral measure when lacking an actual one, and therefore its results might not be comparable to the other methods. Overall, the magnitudes and behaviors of the spectra measures for different methods are very different and may be difficult to compare.

Spectral distribution probability (SD probability) and entropy (SD entropy) tests showed that by our setup and library, the different identification methods indicate quite even certainties in their identification guesses. However, the SID method showed greatest extreme results, which reached the SD probability values of 10^{-2}–10^{-4}, compared to the typical magnitudes of 10^{-2}. This further supports Chang's view of the superiority of the SID method. Therefore, the differentiation tests show the superiority of the SID method, even if the identification-rejection graphs of the methods did not show clear differences between the methods excluding the correla-

tion method.

A major problem in this setup is the drifting behavior of the supercontinuum (SC) light source, which causes difficulties in creating comparable data. A solution has been presented for this problem: the setup could use a detector to analyze the output spectrum, which would be used to eliminate the drifting from the reflection spectra. The downside of this solution is the significantly higher cost and/or more complicated setup architecture: adding to the setup a splitting mirror and another spectrometer would allow continuous supervising of the spectral drifting, but it would significantly raise the cost of the setup. Meanwhile, a precisely moving beam guide would be needed if the setup uses the same spectrometer to measure both the reflection spectra and the raw SC output one after the other. However, the beam guide would be vulnerable to mechanical impacts on the setup, and measuring drift-monitoring spectra would cause the overall measurement to become notably slower. Additionally, this method would prevent the continuous supervision of the SC drifting.

With the current identification software, the calculation time of the identification process effortlessly falls below the scanner step movement time. This gives room for allowing the setup to use a larger material type library and more elaborate identification methods, such as a protocol consisting of multiple phases. If the future versions of the software or convenience require faster download times, the speed can be increased by employing binary file formats.

It was noted that averaging subsequent spectra over time improves the results up to an extent. Most of these tests used averaging that equals either 0.5 or 1 illumination seconds (for the target distances of 14 m and 40 m respectively); the actual measurement time would be at least twice this long, because the device also needs to measure a non-illuminated spectrum. This is a convenient time period for a single-pixel measurement, even though the scan would still be long given that an actual scan may contain hundreds of pixels. When considering the unstable behavior of the SC source, too long single-pixel measurement times endanger the measurement accuracy.

The effect of the noise on the required minimum identification level was also studied in the thesis, but the examination gave no distinct relation between incorrect identifications and too large noise. This is unfortunate, as if the noise level could be used to define the minimum acceptance level for the identification, giving stricter identification acceptance criteria for very noisy data than for clear data. The future work should also address the weighing imbalance inside the material probability clouds. Since more than three-dimensional spaces are challenging to visualize by a human mind, it would be useful to find a numerical method to determine the weighing inside a material probability cloud.

Overall, the results of this thesis recommend using either the SID method with only a few spectral bands or the correlation method with as many spectral bands as possible. Because these two methods behave very differently, combining these two methods and employing their best features could be an interesting choice. However, the SID method showed remarkable superiority only in the differentiation tests, while the identification-rejection graphs did not highlight its performance. This indicates

the uncertain nature of the superiority status of the SID method and/or that the used material libraries hinder the performance of the methods. Generally, ideal minimized material libraries could also improve the performance of the software, which would not only limit the calculation time but also decrease the chance of making wrong material guesses. For example, when identifying targets in the field, specialized libraries could be used according to the season or climate, such as limiting the number of plant spectra during Finnish winter.

In scanning tests, the software was well capable of identifying targets in indoors settings. However, it had problems in outdoors due to temperature shifts that affected the stability of the SC source. Target identification could also be assisted by using intensity data of the scans for detecting reflective (and therefore often man-made) objects. Also RGB maps could be used especially when seeking specified targets; different colors could be used to identify the material probabilities of a single material at each scan pixel.

A major challenge in improving the performance of the setup is addressing the unstable behavior of the SC source. Additionally, compensating for the atmospheric effects with other methods than by the white reference would be ideal, as then the white reference is not required to be brought to the measurement distance for each measurement. Overall, there are still several ways to improve both the physical instrument and its identification software, and it is likely that these improvements could also improve the performance of the material identification.

References

[1] L. Farr, M. Watson, and G. Bishop. Active Spectral Imaging for Target Detection. *3rd EMRS DTC Technical Conference*. Edinburgh, 2006.

[2] M. T. Eismann. *Hyperspectral remote sensing*. SPIE Press, 2012.

[3] V. A. Kovalev and W. E. Eichinger. *Elastic Lidar: Theory, Practice, and Analysis Methods*. John Wiley & Sons, Inc., 2004.

[4] T. Kääriäinen. Long Distance Hyperspectral Lidar for Target Recognition. Master's thesis, Aalto University School of Electrical Engineering, Espoo, 2014.

[5] A. Manninen, T. Kääriäinen, T. Parviainen, S. Buchter, M. Heiliö, and T. Laurila. Long distance active hyperspectral sensing using high-power near-infrared supercontinuum light source. *Optics Express*, 22(6):7172–7, 2014.

[6] C. I. Chang. An information-theoretic approach to spectral variability, similarity, and discrimination for hyperspectral image analysis. *IEEE Transactions on Information Theory* 46, p. 1927–1932, 2000.

[7] R. Pastila. *Ultravioletti- ja lasersäteily*. Säteily- ja ydinturvallisuus. Säteilyturvakeskus, 2009.

[8] H. Ibach and H. Lüth. *Solid-State Physics: An Introduction to Principles of Materials Science*. 4th Edition. Springer, 2009.

[9] C. F. Bohren and D. R. Huffman. *Absorption and scattering of light by small particles*. Wiley, 2008.

[10] M. I. Disney, P. E. Lewis, M. Bouvet, A. Prieto-Blanco, and S. Hancock. Quantifying surface reflectivity for spaceborne lidar via two independent methods. *IEEE Transactions on Geoscience and Remote Sensing*, 47(9):3262–3271, 2009.

[11] F. D. van der Meer, H. M. A. van der Werff, F. J. A. van Ruitenbeek, C. A. Hecker, W. H. Bakker, M. F. Noomeb, *et al.* Multi- and hyperspectral geologic remote sensing: A review. *International Journal of Applied Earth Observation and Geoinformation*, 14 p. 112–128, 2012.

[12] F. A. Kruse, A. B. Lefkoff, J. W. Boardman, K. B. Heidebrecht, A. T. Shapiro, P. J. Barloon, and A. F. H. Goetz. The spectral image-processing system (SIPS) - interactive visualisation and analysis of imaging spectrometer data. *Remote Sensing of Environment* 44, p. 145–163, 1993.

[13] J. M. Dudley, G. Genty, and S. Coen. Supercontinuum generation in photonic crystal fiber. *Reviews of modern physics*, 78(4):1135, 2006.

[14] D. W. Allan. Statistics of Atomic Frequency Standards. *Proceedings of the IEEE* 54, no. 2, p. 221–230, 1966.

[15] M. Kraft, A. Kenda, A. Frank, W. Scherf, A. Heberer, T. Sandner, H. Schenk, and F. Zimmer. Single-detector micro-electro-mechanical scanning grating spectrometer. *Analytical and Bioanalytical Chemistry*, 386(5):1259–1266, Oct 2006.

[16] D. Manolakis, R. Lockwood, T. Cooley, and J. Jacobson. Is there a best hyperspectral detection algorithm?. *Algorithms and Technologies for Multispectral, Hyperspectral, and Ultraspectral Imagery XV.* Edited by S. S. Shen and P. E. Lewis. Orlando, Florida, USA. SPIE. 733402-16. 2009

[17] G. Mountrakis, J. Im, and C. Ogole . Support vector machines in remote sensing: a review. *ISPRS Journal of Photogrammetry and Remote Sensing* 66, p. 247–259, 2011.

[18] F. van der Meer and W. Bakker. Cross correlogram spectral matching: application to surface mineralogical mapping by using AVIRIS data from Cuprite, Nevada. *Remote Sensing of Environment* 61, p. 371–382, 1997.

[19] X. Zhang and P. Li. Lithological mapping from hyperspectral data by improved use of spectral angle mapper. *International Journal of Applied Earth Observation and Geoinformation* 31, p. 95–109, 2014.

[20] J. W. Rouse Jr., R. H. Haas, J. A. Schell, D. W. Deering, and J. C. Harlan. Monitoring the vernal advancement and retrogradation (green wave effect) of natural vegetation. *NASA/GSFC, Final Report, Greenbelt, MD, USA*, p. 1–137. 1974.

[21] V. Henrich. *Index DataBase: A database for remote sensing indices*, 2011. [online] Available from: `http://www.indexdatabase.de/` [Accessed 4 Sep. 2014].

[22] L. S. Rothman, I. E. Gordon, Y. Babikov, A. Barbe, D. Chris Benner, P. F. Bernath, M. Birk, L. Bizzocchi, V. Boudon, L. R. Brown, *et al.* The HITRAN2012 molecular spectroscopic database. / *Journal of Quantitative Spectroscopy & Radiative Transfer*. 130 p. 4–50, 2013.

A Correct identification ratio and rejection ratio

The following identification-rejection graphs (see Chapter 5.1.1) feature the correct identification ratios and rejection ratios for each identification method with two different library choices (non-alpha and alpha) and three different material category types (all categories, subcategories, supercategories). For the graphs, the test spectra were created by averaging 5 subsequent snapshot spectra and non-averaged WR spectra. The black crosses indicate the average correct identification ratios for the configurations with rejection ratios of 0–25%.

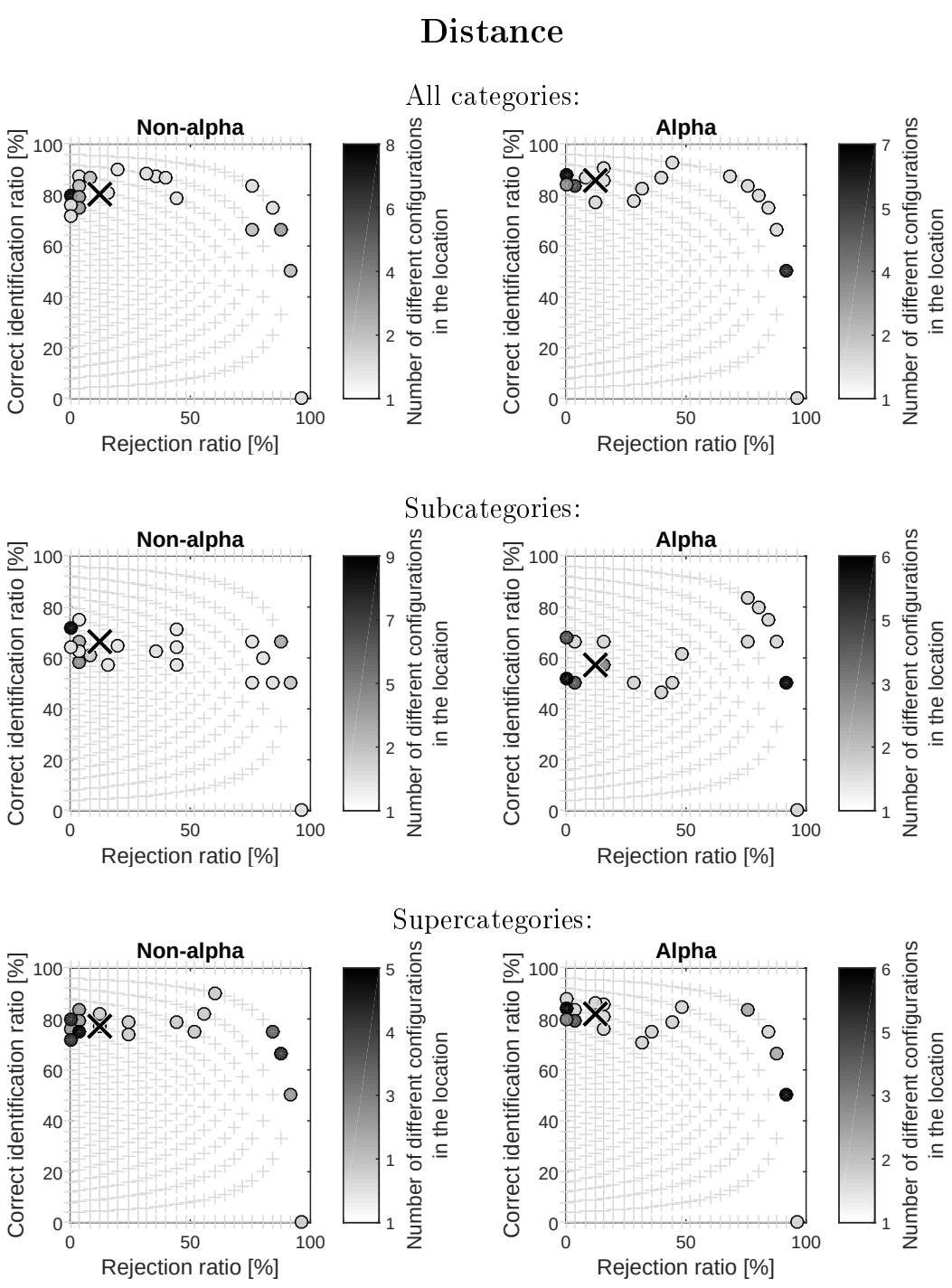


Figure A1: Rejection ratio versus correct identification ratio for the distance method, using all categories (upper row), only subcategories (middle row), or only supercategories (lower row), with the non-alpha library (left) or the alpha library (right).

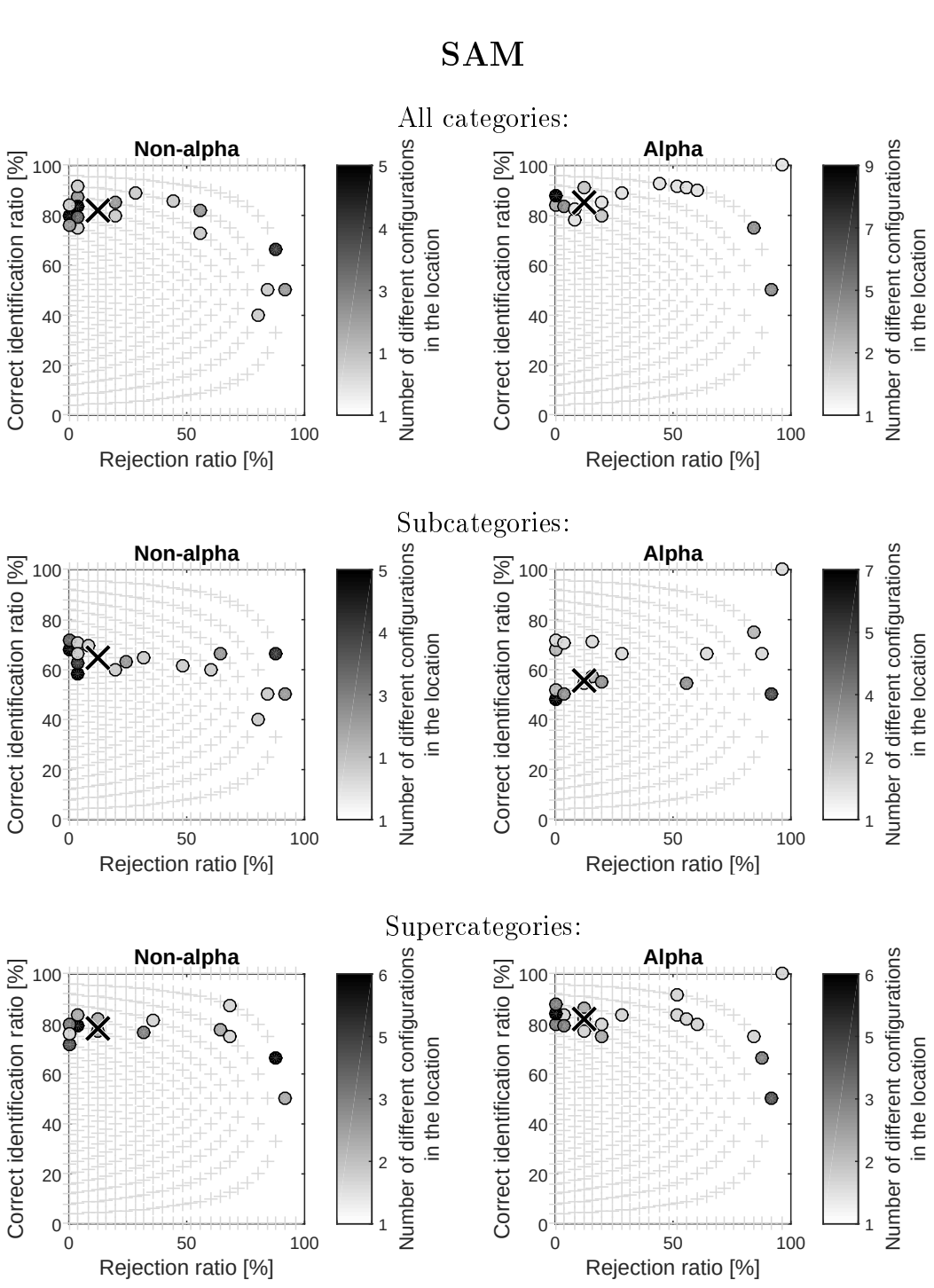


Figure A2: Rejection ratio versus correct identification ratio for the SAM method, using all categories (upper row), only subcategories (middle row), or only supercategories (lower row), with the non-alpha library (left) or the alpha library (right).

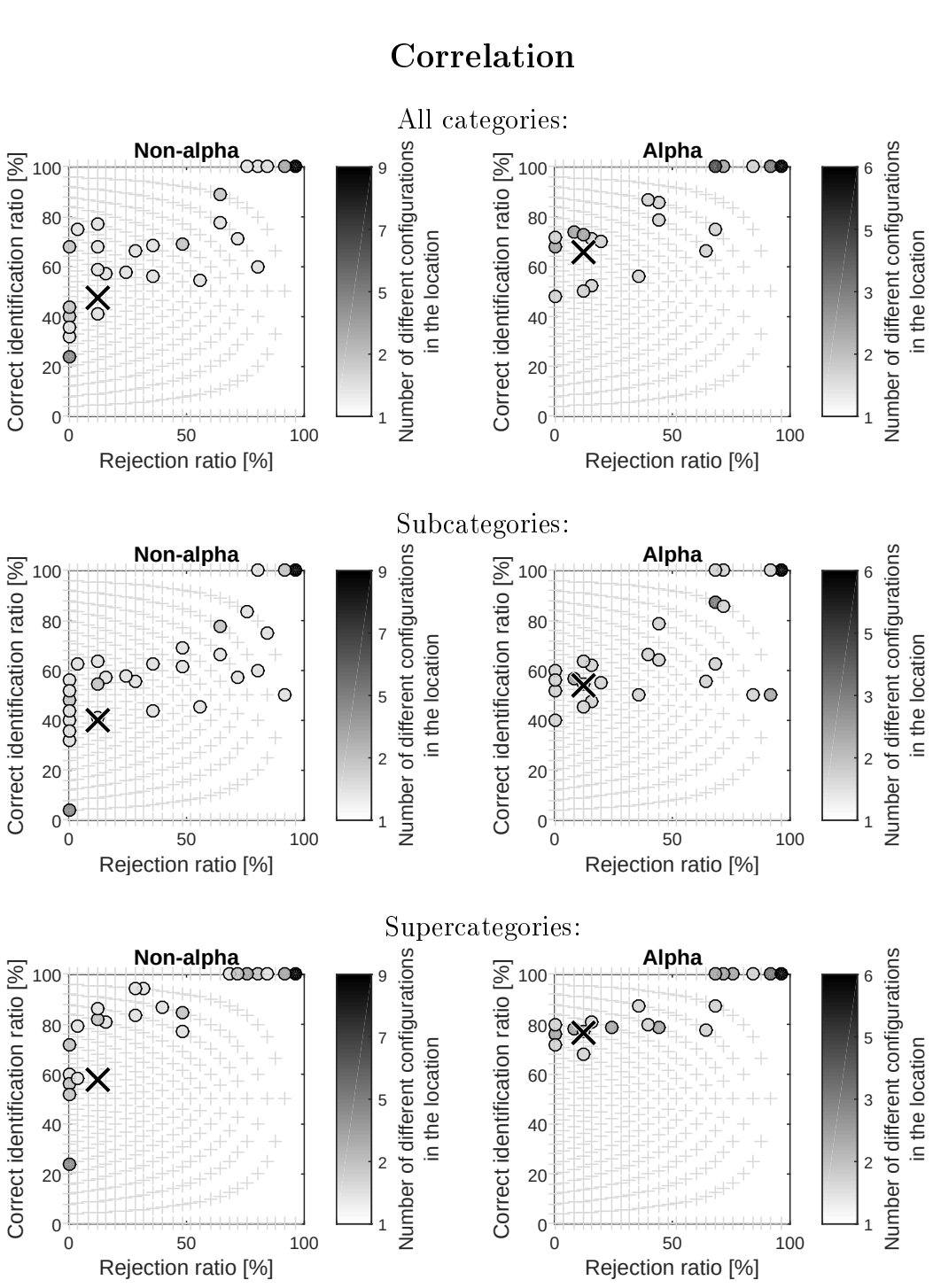


Figure A3: Rejection ratio versus correct identification ratio for the correlation method, using all categories (upper row), only subcategories (middle row), or only supercategories (lower row), with non-alpha library (left) or alpha library (right).

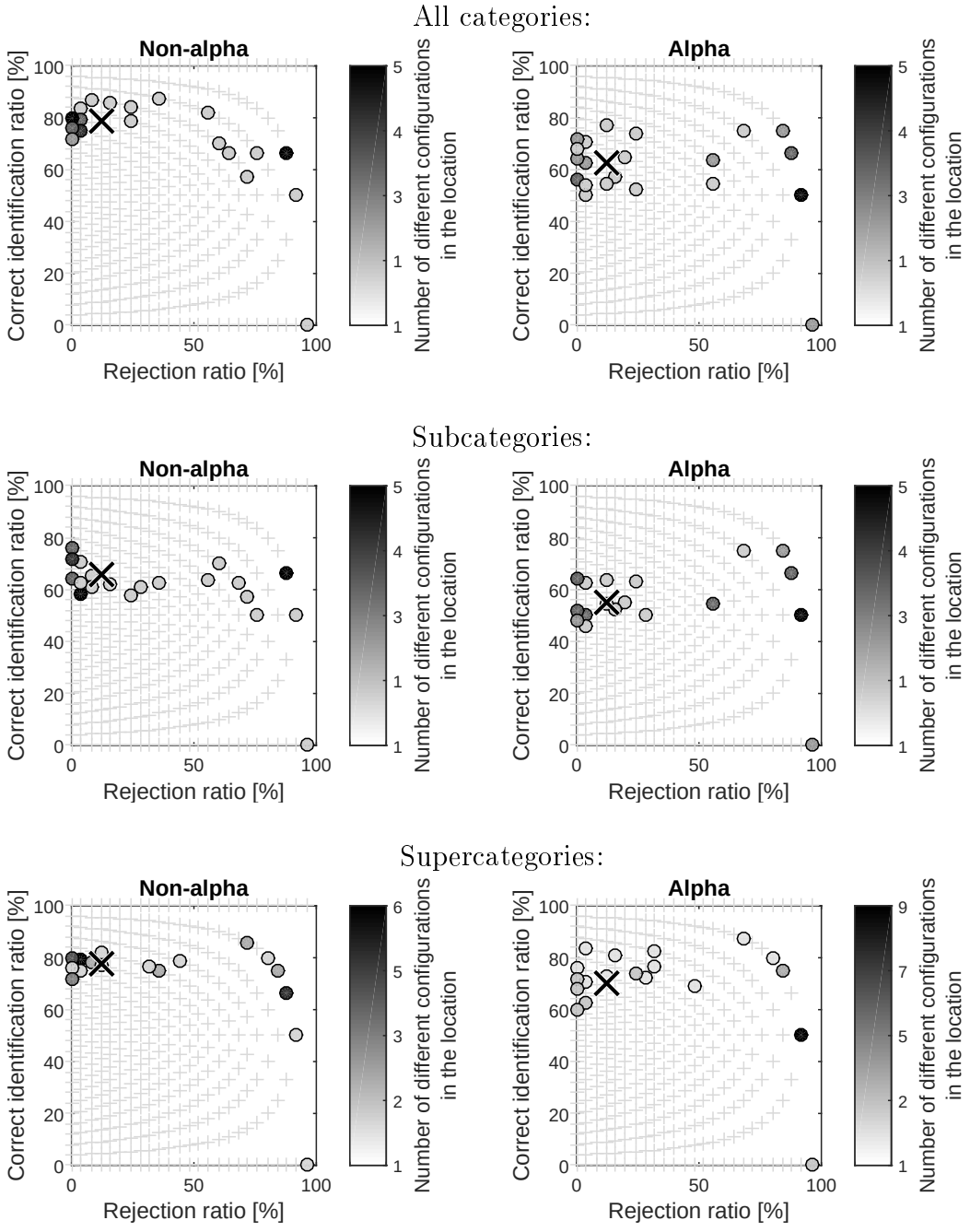


Figure A4: Rejection ratio versus correct identification ratio for the MND method, using all categories (upper row), only subcategories (middle row), or only supercategories (lower row), with non-alpha library (left) or alpha library (right).

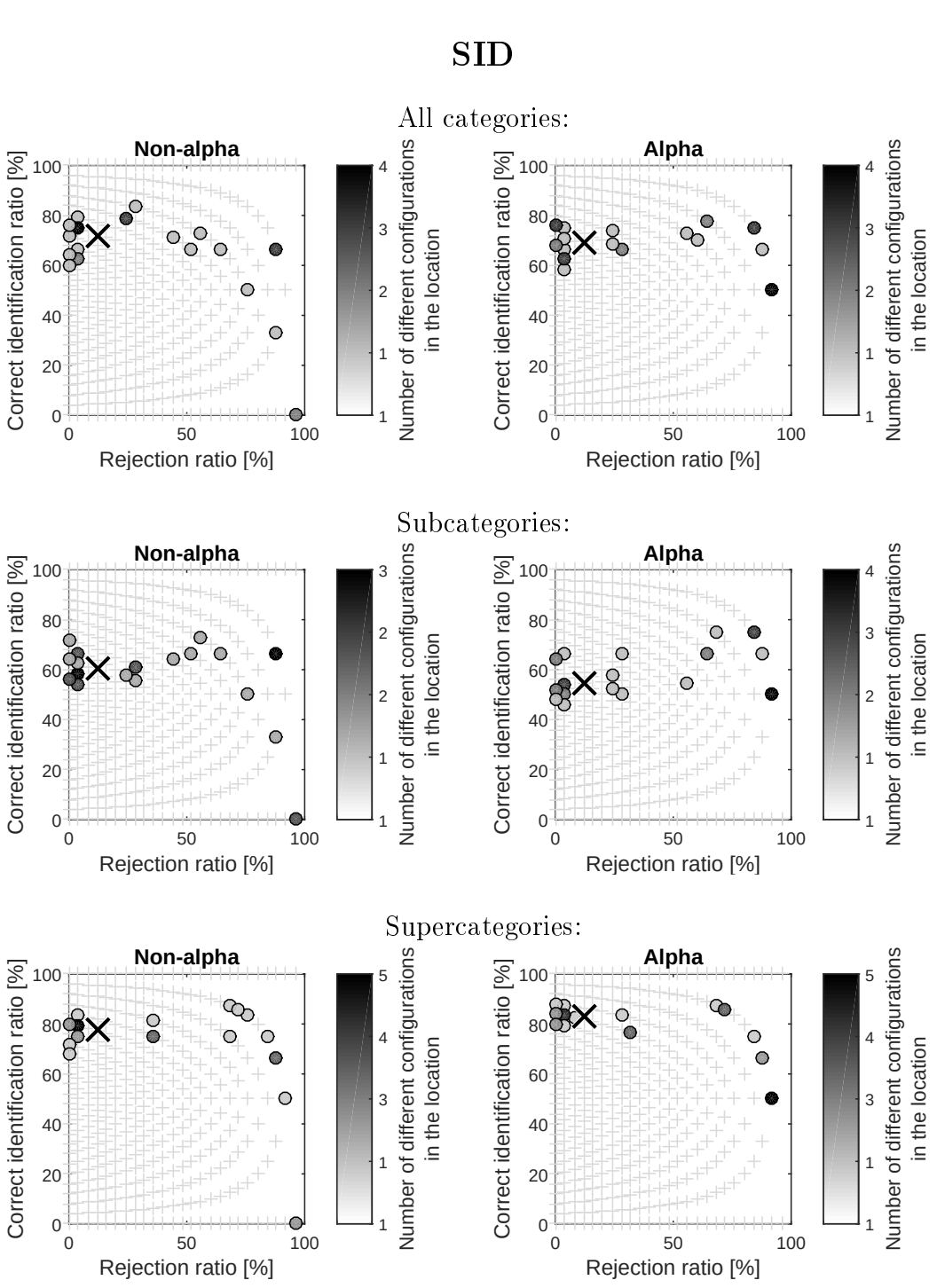


Figure A5: Rejection ratio versus correct identification ratio for the SID method, using all categories (upper row), only subcategories (middle row), or only supercategories (lower row), with the non-alpha library (left) or the alpha library (right).

B Effect of band configurations and variance parameters

The following figures feature the effect of the band division choice and the variance parameters on the correct identification ratio and the rejection ratio. The test spectra were created by averaging 5 subsequent snapshot spectra and non-averaged WR spectra. The greyscale dots express either the correct guess ratio (the upper graph of the graph pairs) or the rejection ratio (the lower graph). Crosses mark configurations that rejected all test spectra, black diamonds mark the identification rate of 100%, and white diamonds mark either identification or rejection value of 0% depending on the graph.

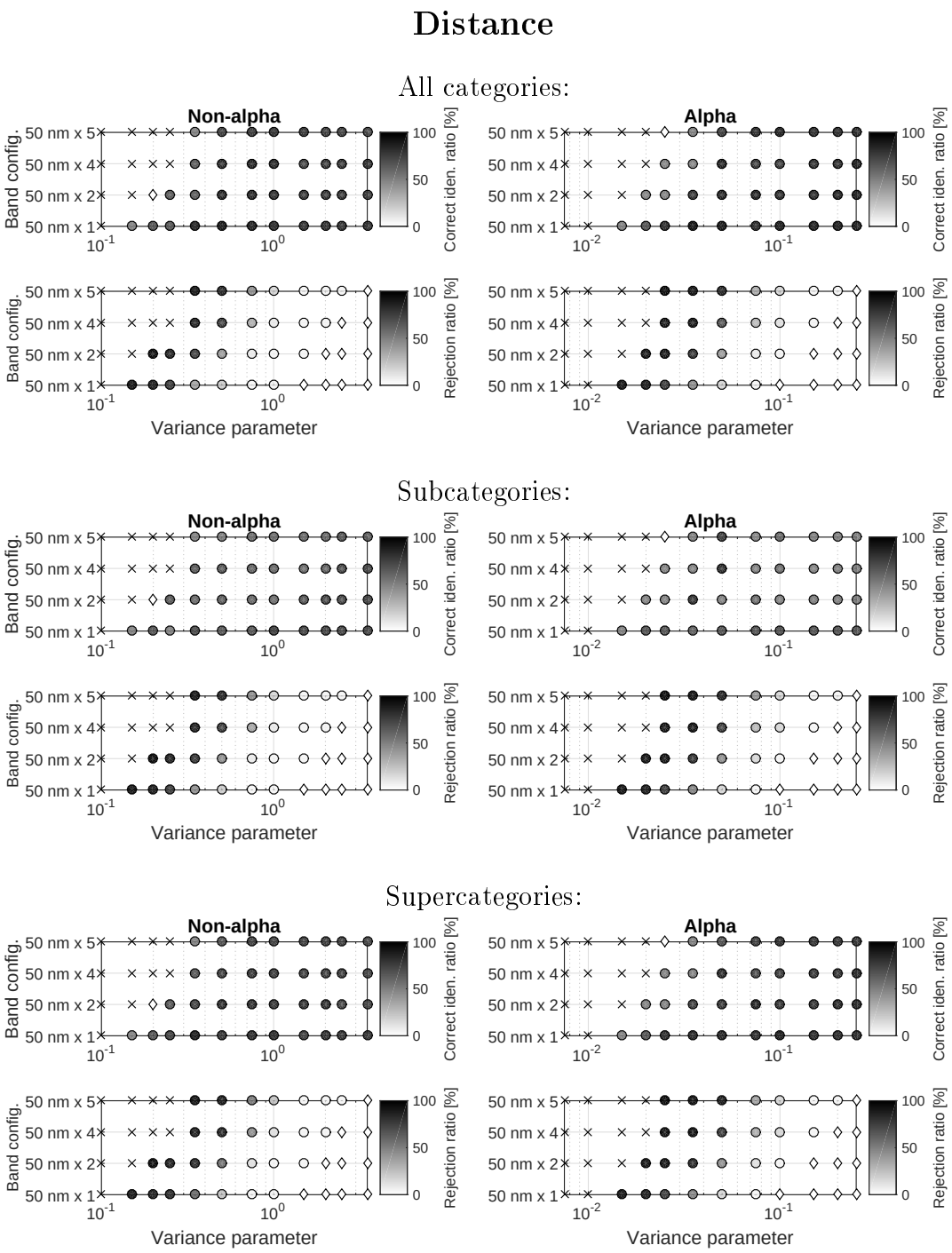


Figure B1: The effect of band choices and variance parameters on the correct identification ratios and rejection ratios for the distance method, using all categories (upper row), only subcategories (middle row), or only supercategories (lower row), with the non-alpha library (left) or the alpha library (right).

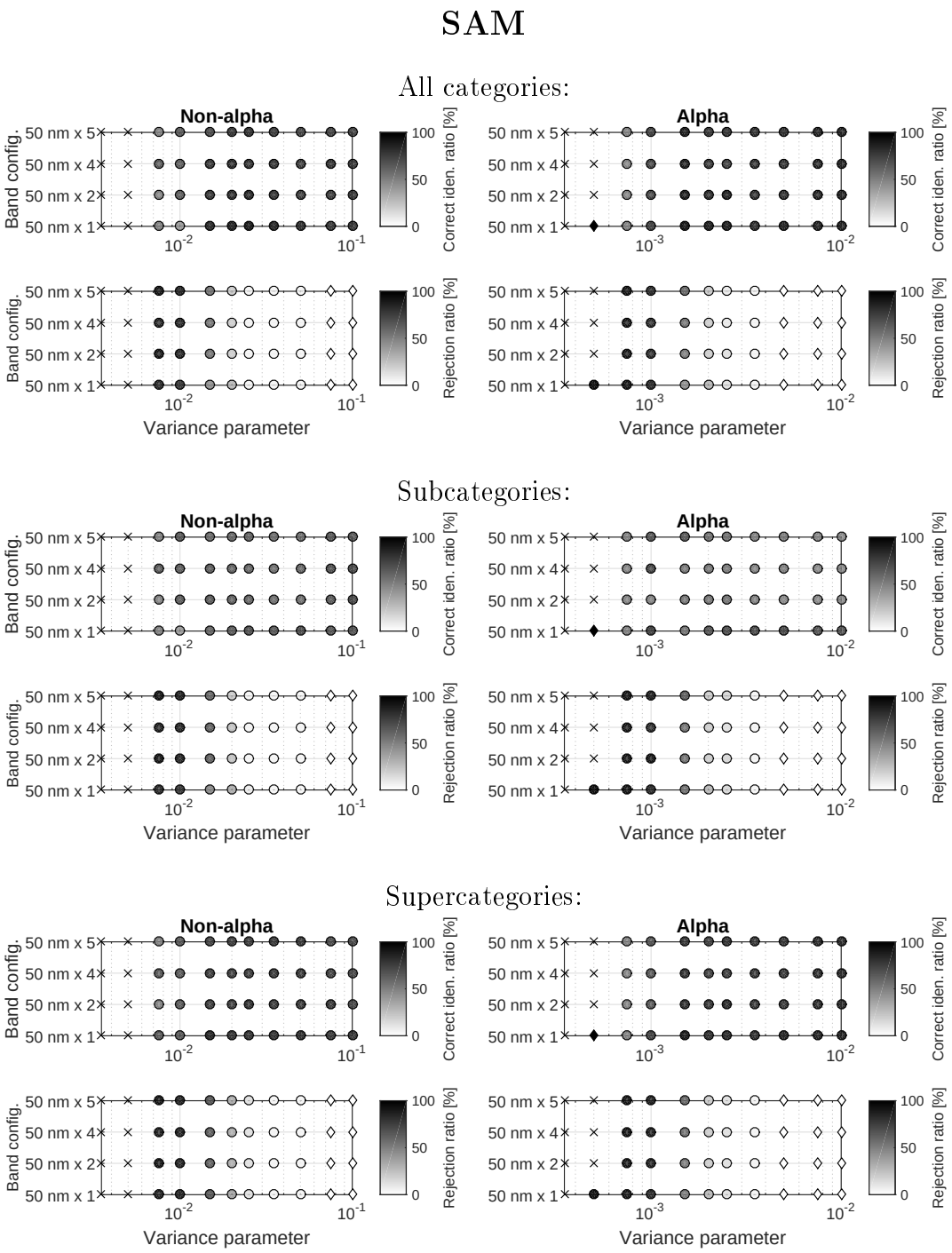


Figure B2: The effect of band choices and variance parameters on the correct identification ratios and rejection ratios for the SAM method, using all categories (upper row), only subcategories (middle row), or only supercategories (lower row), with the non-alpha library (left) or the alpha library (right).

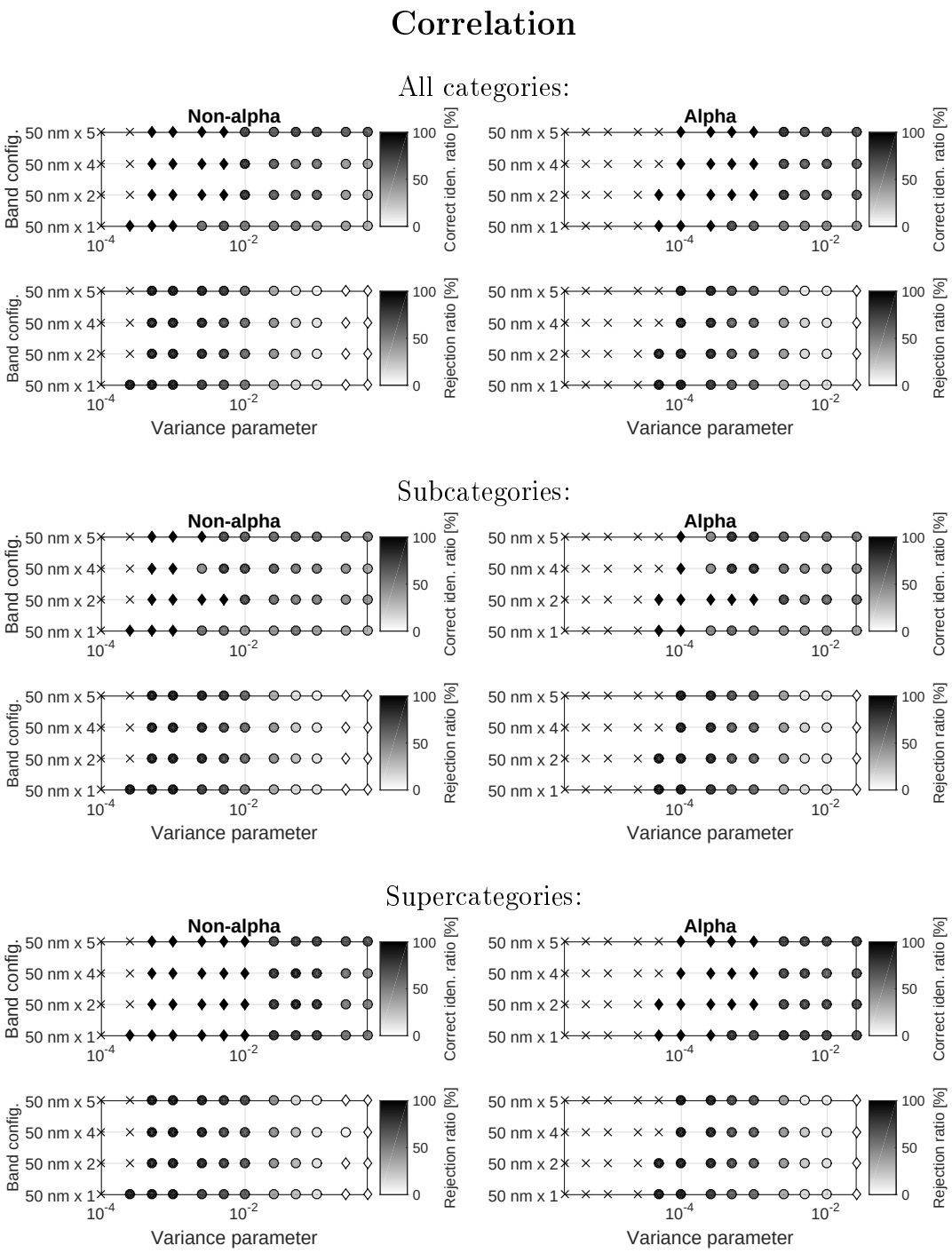


Figure B3: The effect of band choices and variance parameters on the correct identification ratios and rejection ratios for the correlation method, using all categories (upper row), only subcategories (middle row), or only supercategories (lower row), with the non-alpha library (left) or the alpha library (right).

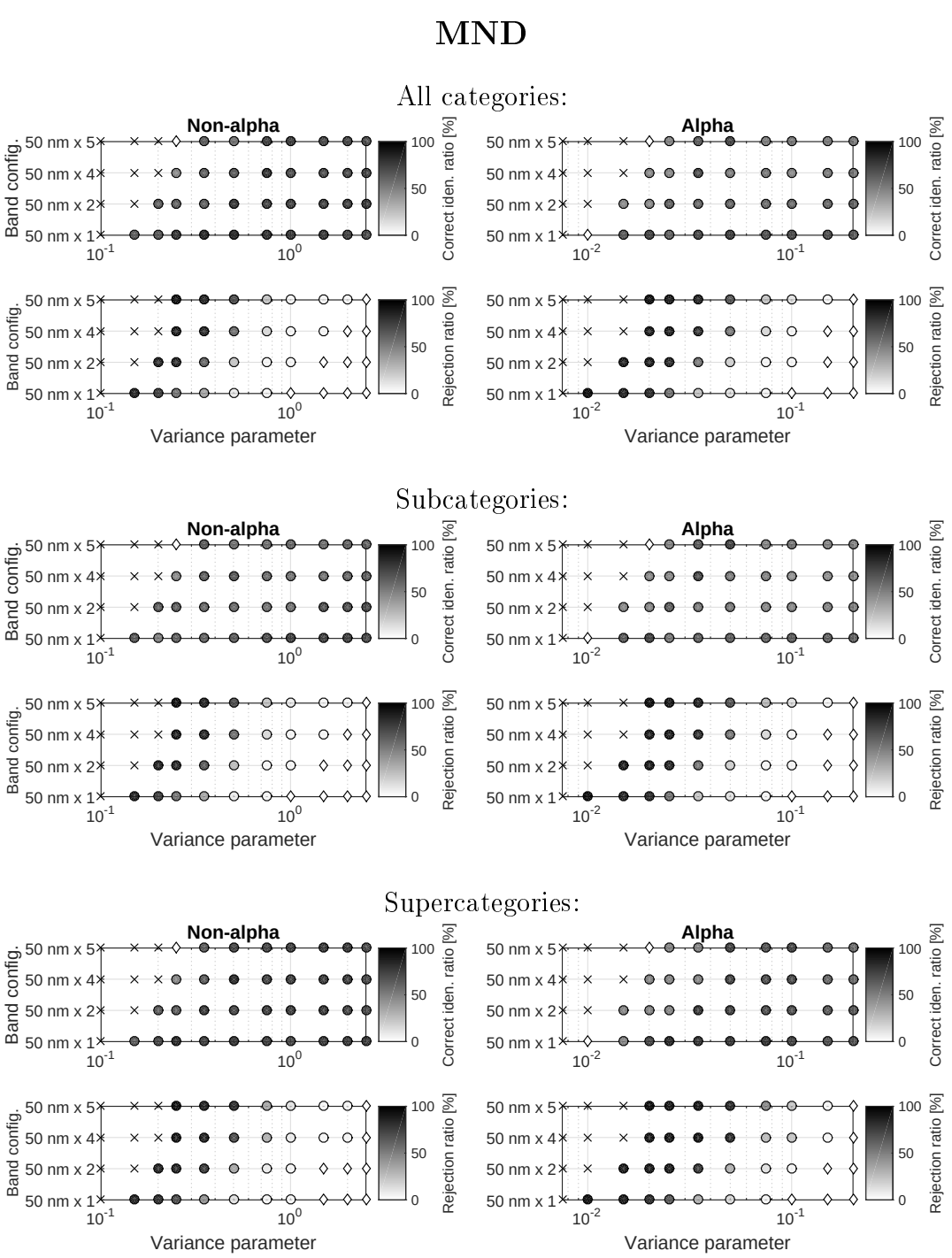


Figure B4: The effect of band choices and variance parameters on the correct identification ratios and rejection ratios for the MND method, using all categories (upper row), only subcategories (middle row), or only supercategories (lower row), with the non-alpha library (left) or the alpha library (right).

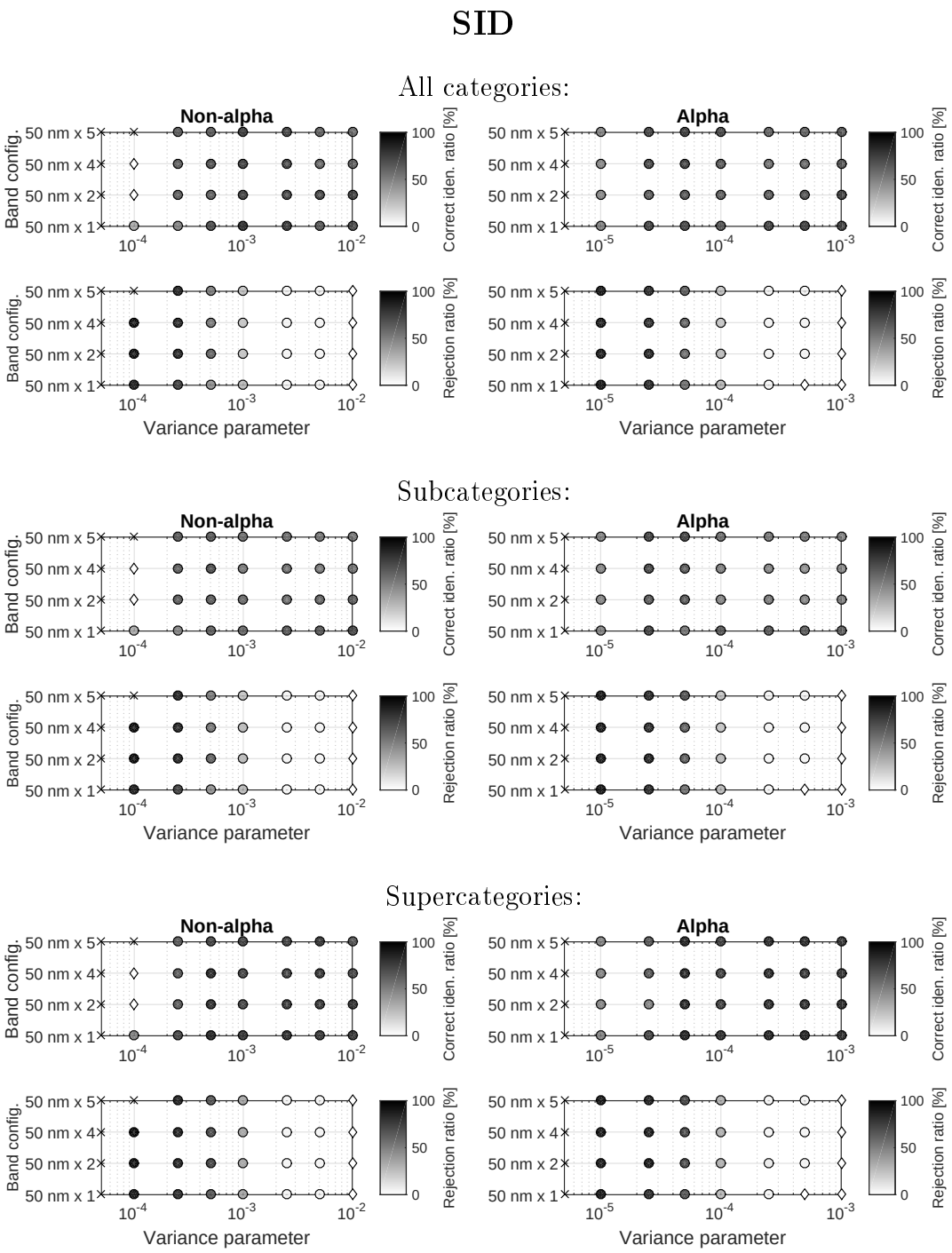


Figure B5: The effect of band choices and variance parameters on the correct identification ratios and rejection ratios for the SID method, using all categories (upper row), only subcategories (middle row), or only supercategories (lower row), with the non-alpha library (left) or the alpha library (right).